

Departamento de Ciências e Tecnologias da Informação

PALbot: a Plug&(Almost)pLay chatbot

André Gilberto Tai

Dissertação submetida como requisito parcial para obtenção do grau de
Mestre em Gestão de Sistemas de Informação

Orientador:
Professor Auxiliar
Ricardo Daniel Santos Faro Marques Ribeiro
ISCTE-IUL

Coorientadora:
Professora Auxiliar
Maria Luísa Torres Ribeiro Marques da Silva Coheur
IST, Universidade de Lisboa

06, 2019

Agradecimentos

Ao meu orientador, Professor Ricardo Ribeiro e à minha coorientadora Professora Luísa Coheur, agradeço toda a sua orientação, transmissão de conhecimento, persistência, conselhos oferecidos e tempo dedicado para desenvolver o melhor trabalho possível.

À minha família pelo apoio nos últimos 2 anos, em especial à minha mãe Julieta, ao meu irmão Filipe e à minha namorada Marta pela motivação, auxílio e tolerância nos momentos mais exaustivos permitindo que continuasse e nunca desistir.

Aos meus amigos, Rodrigo Moura, João Santos, Rui Agostinho, Filipe Almeida, João Machado, Gustavo Ferreira, João Sequeira e Nádía Sousa por todo o vosso apoio e amizade.

Por fim, o meu agradecimento para a equipa Ciberdúvidas e para o grupo L2F pela disponibilização dos dados e pelo suporte.

Resumo

Para a comunicação entre seres humanos e computadores é necessária a identificação dos traços que os distinguem e das características do funcionamento de cada um. Desde a eletrónica, que é desenvolvida através de fenómenos físicos bem delineados, à informática, que assenta em axiomáticas matemáticas, o processo de funcionamento de um computador está livre de traços humanos e é conhecido por fornecer respostas concretas, verdadeiras e lógicas.

Neste projeto foi desenvolvido um sistema de diálogo, denominado por PALbot com o objetivo de melhorar a acessibilidade na procura de informação para os utilizadores da plataforma do Ciberdúvidas. O PALbot é um programa que simula uma conversa humana, esclarecendo dúvidas relativamente à Língua Portuguesa, na forma de diálogo com o utilizador. Após desenvolvimento do *chatbot*, foi realizado um inquérito de forma a avaliar a ferramenta desenvolvida. Os resultados do inquérito revelam 65,54% de respostas corretas com a base de conhecimento da plataforma do Ciberdúvidas inserida no sistema de conversação.

Adicionalmente, elaborou-se um conjunto de testes de forma a avaliar os atributos de qualidade, nomeadamente eficácia e eficiência, concluindo que 65,28% dos inquiridos concordam que o PALbot facilita quase sempre ou sempre o acesso à informação.

Em suma, o chatbot desenvolvido durante a presente dissertação apresenta potencial para melhorar a acessibilidade dos utilizadores na procura de informação bem como para auxiliar a plataforma Ciberdúvidas melhorando o esclarecimento de dúvidas relativas à Língua Portuguesa.

Palavras-Chave: Agente de Conversação; Inteligência Artificial; Usabilidade; Sistema de diálogo.

Abstract

A good interaction between humans and computers is built on top of each other's characteristics, functionality and behavior. From electronics, developed through actual physical phenomena, to informatics, built on mathematical axioms, the functionality of a computer is free of human trace. While computers are known for providing correct, true and logic answers, humans spend a vast amount of time on an informal dialog with almost no purpose.

In this project, a dialog system named PALbot was developed to help Ciberdúvidas platform users access information. This program simulates a human conversation answering Portuguese linguistic questions to the user. After developing the chatbot, an inquire was made to evaluate the tool in question showing the program gave a correct answer 65.54% of times with the knowledge base of Ciberdúvidas platform.

Additionally, another pool of tests was made to evaluate quality attributes such as efficiency and efficacy, in which was concluded that 65.28% of the enquired agree that PALbot always and almost always makes it easy to access to the information required.

The chatbot developed during the current dissertation shows high potential to improve user's accessibility to information as well as to help Ciberdúvidas platform answering Portuguese linguistic questions.

Keywords: Chatbot; Artificial Intelligence; Usability; Dialogue System.

Índice

Agradecimentos	i
Resumo	iii
Abstract	v
Índice	vii
Índice de Tabelas	ix
Índice de Figuras	xi
Lista de Abreviaturas e Siglas	xiii
Capítulo 1 – Introdução	1
1.1. Enquadramento do tema.....	1
1.2. Motivação e relevância do tema.....	1
1.3. Questões e objetivos de investigação	2
1.4. Abordagem metodológica	3
1.5. Estrutura e organização da dissertação.....	3
Capítulo 2 – Revisão da Literatura.....	5
2.1. Introdução.....	5
2.2. Recuperação de informação Vs. Geração da língua natural.....	8
2.3. <i>Frameworks</i> para desenvolvimento de <i>chatbots</i>	8
2.3.1. Microsoft	8
2.3.2. Messenger.....	10
2.3.3. Lex.....	10
2.3.4. Watson.....	11
2.3.5. Diagramflow	13
2.3.6. Rasa	14
2.4. <i>Chatbots</i> na língua portuguesa	15
2.4.1. O agente conversacional para o Palácio de Monserrate	15
2.4.2. O agente conversacional baseado em legendas de filmes	16
2.5. <i>Chatbots</i> na língua estrangeira	18
2.5.1. O agente conversacional do Exército	18
2.5.2. O agente conversacional Mágico de Oz	20
2.5.3. O agente conversacional Max.....	21
2.6. Elementos para a avaliação de um <i>chatbot</i>	22
2.7. Discussão.....	24
Capítulo 3 – PALbot.....	27
3.1. Introdução.....	27
3.2. Processamento de dados.....	28

3.3.	Versão baseada na ALICE	30
3.3.1.	Visão Geral	30
3.3.2.	Marcações <i>Artificial Intelligence Markup Language</i>	32
3.3.3.	Redução simbólica.....	35
3.3.4.	Dividir e conquistar	37
3.3.5.	Redução de sinónimos	38
3.3.6.	<i>Set</i> e <i>Get</i>	38
3.3.7.	Outros caracteres <i>wildcard</i>	39
3.3.8.	Interações fora do domínio	41
3.4.	Versão baseada no <i>Say Something Smart</i>	41
3.4.1.	Funcionamento do <i>Say Something Smart</i>	42
3.4.2.	Configurações do <i>Say Something Smart</i>	43
3.5.	Comparação entre as <i>frameworks</i> consideradas.....	44
3.6.	Controlo da base de conhecimento.....	47
3.7.	Aprendizagem	49
Capítulo 4	– Avaliação e testes	51
4.1.	Apresentação do estudo e resultados obtidos	51
4.1.1.	Caracterização da amostra	53
4.1.2.	Avaliação em ambiente controlado	54
4.1.3.	Avaliação em ambiente livre	58
4.1.4.	Análise de relevância do PALbot	61
4.2.	Avaliação do PALbot e do Ciberdúvidas	63
4.3.	Discussão.....	67
Capítulo 5	– Conclusão	69
5.1.	Principais conclusões	69
5.2.	Contributos para a comunidade científica e empresarial	70
5.3.	Limitações do estudo.....	70
5.4.	Propostas de investigação futura	70
Bibliografia		71
Anexos		73

Índice de Tabelas

Tabela 1 – Frameworks representativos de agentes conversacionais.....	25
Tabela 2 – Comparação entre a ALICE e o SSS de acordo com as fases e subfases.....	45
Tabela 3 – Tabela com informações caso as respostas sejam incorretas.....	48
Tabela 4 – Tabela com informações caso o primeiro candidato seja resposta correta...	49
Tabela 5 – Tabela com informações caso o terceiro candidato seja correto.	49
Tabela 6 – Tabela com informações mais complementares.	49
Tabela 7 – Controlo da aprendizagem do PALbot.	50
Tabela 8 – Caracterização dos objetivos das perguntas elaboradas no inquérito.	52
Tabela 9 – Caracterização do intervalo de idades com as gerações.	53
Tabela 10 – Caracterização sociodemográfica dos inquiridos.	53
Tabela 11 – Resultados das perguntas solicitadas mediante as respostas corretas e incorretas.	63
Tabela 12 – Resultados da localização das respostas corretas nos candidatos no PALbot.	63
Tabela 13 – Resultados da localização das respostas corretas nos candidatos no motor de pesquisa do Ciberdúvidas	64
Tabela 14 – Comparação dos resultados de pesquisa entre Ciberdúvidas e PALbot.	66
Tabela 15 – Comparação dos resultados de pesquisa entre Ciberdúvidas e PALbot.	66

Índice de Figuras

Figura 1 – Resultados dos indicadores obtidos no estudo piloto do ensino secundário organizado por gêneros. Todas as questões são baseadas em uma escala que varia de 0 (significado de “not at all”) a 10 (significado de “very much”).	7
Figura 2 – Caracterização das opiniões dos agentes de conversação Chatbot e ALICE quanto aos campos Interesting, Fun e Engaged.	7
Figura 3 – Fluxograma de funcionamento de um bot da Microsoft.	9
Figura 4 – Diagrama de funcionamento da Watson.	12
Figura 5 – Exemplo de um Dialogflow's da Google com solicitação webhook.	13
Figura 6 – Diagrama de funcionamento do Rasa.	15
Figura 7 – Atributos de Qualidade de Chatbots e Agentes de Conversação.	23
Figura 8 – Formato das perguntas e respostas do SSS.	28
Figura 9 – Diagrama de atividades do PALbot, versão SSS.	29
Figura 10 – Dados de ciberdúvidas em Excel.	29
Figura 11 – Diagrama de atividades do PALbot, versão AIML.	30
Figura 12 – Exemplo de correspondência de padrões no AIML.	30
Figura 13 – Diagrama de classes da ALICE.	32
Figura 14 – Marcação AIML para saudação.	33
Figura 15 – Diálogo entre utilizador e ISCTE-IUL.	33
Figura 16 – Marcação AIML para interação de resposta.	34
Figura 17 – Marcação AIML para respostas aleatórias mediante o topic.	34
Figura 18 – Múltiplas marcações AIML.	35
Figura 19 – Diagrama de funcionamento da marcação AIML.	36
Figura 20 – Múltiplas marcações AIML para saudação.	36
Figura 21 – Marcação AIML para dois tipos de entrada.	37
Figura 22 – Diagrama de funcionamento da marcação AIML.	37
Figura 23 – Múltiplas entradas AIML de saudação para respostas aleatórias.	38
Figura 24 – Definição de variável em AIML.	39
Figura 25 – Marcação AIML usando variável.	39
Figura 26 – Marcação AIML com wildcard _.	39
Figura 27 – Marcação AIML com wildcard #.	40
Figura 28 – Prioridade de wildcards.	40
Figura 29 – Diferença de wildcard \$ e _ na marcação AIML.	40
Figura 30 – Respostas default na marcação AIML.	41
Figura 31 – Diagrama de classes do SSS.	41
Figura 32 – Critério na seleção de resposta.	42
Figura 33 – Campo de alteração de linguagem no ficheiro config.txt do SSS.	43
Figura 34 – Configuração de normalizers e stopwords no SSS.	44
Figura 35 – Respostas de fora de domínio.	44
Figura 36 – Diagrama de atividades sobre a apreciação das respostas pelo utilizador.	47
Figura 37 – Diálogo entre o utilizador e o PALbot para a confirmação da viabilidade da informação.	47
Figura 38 – Diálogo entre o utilizador e o PALbot caso a resposta seja incorreta.	48
Figura 39 – Diálogo entre o utilizador e o PALbot caso a resposta seja correta.	48
Figura 40 – Conjunto de pergunta e resposta da base de conhecimento.	50
Figura 41 – Diálogo entre o PALbot e o utilizador.	50
Figura 42 – Indicadores estatísticos da avaliação geral do PALbot e do Ciberdúvidas em ambiente controlado.	54

Figura 43 – Indicadores estatísticos das gerações em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente controlado.....	55
Figura 44 – Indicadores estatísticos das habilitações literárias em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente controlado.	56
Figura 45 – Indicadores estatísticos das áreas de formação em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente controlado.	57
Figura 46 – Indicadores estatísticos da avaliação geral do PALbot e do Ciberdúvidas em ambiente livre.	58
Figura 47 – Indicadores estatísticos das gerações em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente livre.....	59
Figura 48 – Indicadores estatísticos das habilitações literárias em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente livre.	60
Figura 49 – Indicadores estatísticos das áreas de formação em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente livre.	61
Figura 50 – Indicadores estatísticos do histórico de iterações com chatbots.	61
Figura 51 – Indicadores estatísticos da facilidade de acesso à informação fornecida pelo chatbot.	62
Figura 52 – Indicadores estatísticos do histórico de iterações com a facilidade do acesso à informação fornecida pelo chatbot.	62
Figura 53 – Diálogo entre o utilizador e o PALbot.	65
Figura 54 – Pesquisa na plataforma Ciberdúvidas (https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=substantivo+coletivo+da+palavra+habitantes).	65

Lista de Abreviaturas e Siglas

UML – Unified Modeling Language
SPSS – Statistical Package for the Social Sciences
ALICE – Artificial Linguistic Internet Computer Entity
NLP – Natural Language Processing
RNN – Recurrent Neural Networks
Azure AD – Azure Active Directory
LUIS – Language Understanding Intelligent Service
APIs – Application Programming Interfaces
NLU – Natural Language Understanding
ASR – Automatic Speech Recognition
HTTP – HyperText Transfer Protocol
L2F – Laboratório de Língua Falada
POS *tagging* – Part-Of-Speech *Tagging*
OOD – Out of Domain
SSS-Say Something Smart
SGT – Sergeant
ECA – Embodied Conversational Agents
B2C – *Business to Customer*
CTT – Correios de Portugal
AIML – Artificial Intelligence Markup Language

Capítulo 1 – Introdução

1.1. Enquadramento do tema

Um *chatbot* é um programa que simula conversas humanas utilizando inteligência artificial. O agente de conversação é projetado para proporcionar um diálogo numa linguagem utilizada naturalmente pelos seres humanos, permitindo uma disponibilidade em tempo integral e acesso à distância (Pereira, Coheur, Fialho & Ribeiro, 2016).

Usualmente os *chatbots* são usados em sites com o objetivo de auxiliar os utilizadores a encontrar produtos e esclarecer dúvidas. Neste sentido, pretende-se uma melhoria na procura de informação por parte dos utilizadores, particularmente através do *chatbot*. Para o desenvolvimento deste *chatbot* em particular será utilizada a informação do Ciberdúvidas como base de conhecimento, uma plataforma de informação, esclarecimento de dúvidas, debate e promoção da língua portuguesa que está alojado nos servidores do ISCTE-IUL.

Em síntese, este projeto pretende desenvolver um *chatbot* denominado por PALbot para a plataforma Ciberdúvidas.

1.2. Motivação e relevância do tema

Segundo Abílio Oliveira (2011, p.72), em O Desafio da Vida, “O termo educar tem a sua origem etimológica no verbo latino *educare*, derivado de *educere*, ou eduzir, e significa conduzir para o exterior, despertar no Homem os elementos positivos que nele se achavam dormentes, como sejam a verdade, a justiça, o amor, a tolerância ou a solidariedade.”.

Este projeto foi realizado com o objetivo de desenvolver um sistema que permita auxiliar os utilizadores da plataforma do Ciberdúvidas a encontrar a informação relativamente à Língua Portuguesa de forma simplificada.

Salientando o sucesso dos *chatbots* em diversos campos (Pereira *et al.*, 2016), revela-se importante verificar se o mesmo se sucede num contexto educacional, em particular relacionado com a Língua Portuguesa.

A principal motivação para a escolha deste tema surgiu da necessidade em assegurar qualidade e eficácia na informação que os utilizadores conseguem obter a partir da plataforma do Ciberdúvidas. A relevância deste projeto acresce ainda pela perspectiva

educacional do mesmo, numa sociedade cada vez mais inclinada para as novas tendências da tecnologia. Neste sentido, considera-se que esta constitui um suporte essencial ao desenvolvimento deste projeto.

1.3. Questões e objetivos de investigação

A presente dissertação propõe-se responder à seguinte questão: Como desenvolver um agente conversacional capaz de melhorar a acessibilidade na procura de informação para os utilizadores na plataforma do Ciberdúvidas?

A questão de partida levou-nos a definir os seguintes objetivos:

- Analisar e desenvolver um *chatbot*. Será importante analisar quais são os parâmetros e métodos para desenvolver esta ferramenta. Os resultados da análise gerados serão posteriormente utilizados para a criação de um *chatbot*. Por fim, os resultados da análise serão representados em diagramas de *Unified Modeling Language*¹ (UML), de forma a apresentar componentes importantes da arquitetura do sistema (Nunes & O’Neil, 2004);

- Compreender o impacto da implementação do *chatbot* na plataforma, através da realização de um inquérito aos utilizadores. Os resultados obtidos nesta fase serão analisados pela ferramenta *Statistical Package for the Social Sciences*² (SPSS) para compreender a influência na acessibilidade na procura de informação pelos utilizadores tanto no PALbot como na plataforma Ciberdúvidas.

O principal objetivo da presente dissertação é o desenvolvimento e preparação de um *chatbot* para a plataforma Ciberdúvidas. Assim, revela-se necessário testá-lo de forma a garantir que a ferramenta em questão vai ao encontro dos objetivos definidos. A avaliação do PALbot pelos utilizadores irá ter em conta a correção, eficácia e eficiência percebida das respostas fornecidas pelo *chatbot*.

¹ Linguagem para elaboração de estrutura de projetos de *software*.

² *Software* de estatística produzida pela *International Business Machines*(IBM).

1.4. Abordagem metodológica

De forma a cumprir os objetivos supracitados, o desenvolvimento do trabalho iniciou-se com uma revisão da literatura, da qual se obteve informação relativa aos *chatbots*, nomeadamente:

- Uma breve história do seu surgimento;
- Definições, tipologias e parâmetros para a conceção de *chatbots* existentes;
- Metodologias de comparação de desenvolvimento de *chatbots*.

De seguida, procede-se à análise, desenvolvimento do PALbot, incluindo o processamento dos dados do *corpus*³ enquadrado no Ciberdúvidas.

Por último, irá ser elaborado um inquérito para testar a funcionalidade do PALbot, e através da análise dos seus resultados avaliar se esta ferramenta vai ao encontro dos seus objetivos. Adicionalmente, irá ser realizado um conjunto de testes para verificar se o PALbot é um sistema de diálogo eficiente e eficaz.

1.5. Estrutura e organização da dissertação

A presente dissertação é composta por cinco capítulos.

O primeiro capítulo apresenta o enquadramento teórico, o tema de investigação e os objetivos da mesma, bem como uma descrição sucinta da estrutura do trabalho.

O segundo capítulo reflete a revisão da literatura, onde se apresentam as várias abordagens de desenvolvimento de *chatbots* em diferentes áreas.

O terceiro capítulo consiste na apresentação do projeto, onde são explicitados os objetivos, o levantamento de requisitos, as ferramentas e os parâmetros que são fulcrais para o sucesso do projeto.

O quarto capítulo é dedicado à avaliação do PALbot e ao processo metodológico utilizado aquando da recolha e análise de dados.

O quinto capítulo prende-se com as principais conclusões do estudo realizado, assim como recomendações, limitações e orientações para possíveis trabalhos futuros.

³ Base de conhecimento constituído por pares de interação de perguntas e respostas.

Capítulo 2 – Revisão da Literatura

Uma interface conversacional ou apenas *chatbot* funciona através da simulação de um diálogo entre o Homem e a máquina (Dale, 2016).

Neste capítulo, começar-se-á por apresentar as *frameworks*⁴ que existem para desenvolver *chatbots* e sistemas que foram desenvolvidos para situações específicas em Língua Portuguesa.

Relativamente ao desenvolvimento de *chatbots* a partir de *frameworks*, as oito plataformas apresentadas são Pandorabots, Microsoft, Messenger, Amazon, IBM Watson, Google Diagramflow e Rasa. Os *chatbots*, que não foram desenvolvidos em plataformas, são o Edgar – o Mordomo de Monserrate (Fialho *et al.*, 2013), o Filipe – desenvolvido pelo INESC-ID (Ameixa, Coheur, Fialho & Quaresma, 2014), o *Sergeant Blackwell* (Leuski, Patel & Traum, 2006), o ChatWoz (Fialho & Coheur, 2015) e o agente Max (Jung & Kopp, 2003).

2.1. Introdução

O primeiro *chatbot* surgiu em público em 1966 sob a aparência de um psicoterapeuta chamado Eliza. O Eliza desenvolvido por Joseph Weizenbaum foi o primeiro programa que conseguiu estabelecer uma conversa com o Homem. O modelo de conversação do Eliza foi baseado na reformulação das frases de entrada, quando estas emparelham com um conjunto de regras pré-definidas. O sucesso deste *chatbot* na altura, deveu-se ao facto de muitos dos utilizadores acreditarem que conversavam com outra pessoa, o que deu origem ao atualmente conhecido *Eliza effect*.

O *Artificial Linguistic Internet Computer Entity* (ALICE) foi inventado em 1995 por Richard Wallace. Considerado uma versão mais moderna do Eliza, por ser baseado na correspondência de padrões, difere do mesmo uma vez que se tenta comportar como um ser humano em geral ao invés de desempenhar um papel específico. Inicialmente foi desenvolvido com o objetivo de manter uma conversa com pessoas o maior tempo possível sem que as mesmas se apercebessem que estavam a falar com uma máquina. Associado à ALICE existe uma coleção de recursos que têm sido amplamente utilizados

⁴ *Software* que une códigos de programação.

pela comunidade de *chatbots*, incluindo o serviço de alojamento na plataforma Pandorabots⁵, que representa a maior comunidade *chatbot* na Internet.

Os *chatbots* variam de mais simples, baseados em modelos de correspondência de padrões, explorando grandes armazenamentos de pequenas respostas preparadas, para arquiteturas mais complexas, baseadas nalgum tipo de processo de aprendizagem

Atualmente, existem muitas plataformas para ajudar a desenvolver *chatbots*, e o número de novos *chatbots* continua a aumentar a um ritmo veloz. O serviço de alojamento no Pandorabots declara ter mais de 225.000 *botmasters* — pessoas encarregadas de criar ou de manter um *chatbot*, sendo que já foram desenvolvidos mais de 250.000 *chatbots* no que resultou em mais de três milhares de milhões de interações (Pereira *et al.*,2016).

No entanto, em Portugal não se verifica um acompanhamento deste crescimento no sector das Tecnologias de Informação, justificado por Patrícia Fernandes, diretora de Marketing e Comunicação da Microsoft em Portugal por um número insuficiente de candidatos que possuam as competências necessárias face a um número bastante maior de vagas por preencher, contrariando as estatísticas e tornando necessário criar soluções para colmatar a falta de candidatos nesta mesma área⁶.

O mesmo acontece em outros países, como é o caso da Argentina. De acordo com Benotti, Martínez & Schapachnik, (2014) As universidades argentinas formam cerca de 3.500 estudantes da área de ciência computacional por ano enquanto que na área de direito formam 10.000 e 15.000 em economia. A indústria nacional argentina necessita de mais estudantes formados na área ciência computacional e parte desse problema provém da ausência de aulas de programação no ensino secundário. A ferramenta de *software*⁷ educacional desenvolvida pela Universidade Nacional de Córdoba com objetivo de incentivar os alunos do ensino secundário a aprender conceitos básicos da ciência computacional é designado por Chatbot. De forma a analisar o envolvimento dos alunos com a área de ciência computacional, foi realizado um estudo, que consiste em um conjunto de lições direcionado para uma turma de uma escola secundária Argentina com o objetivo de reunir informações mais detalhadas sobre o comportamento dos alunos usando as ferramentas ALICE e Chatbot. Após o curso, foi aplicado um inquérito aos

⁵ <https://www.pandorabots.com/>

⁶ <https://expresso.pt/economia/2015-10-18-Tecnologia-O-sector-onde-o-que-falta-sao-os-candidatos-#gs.eskdir>

⁷ Suporte lógico.

alunos com perguntas que estão presentes na Figura 1 (Benotti, Martínez & Schapachnik, 2014).

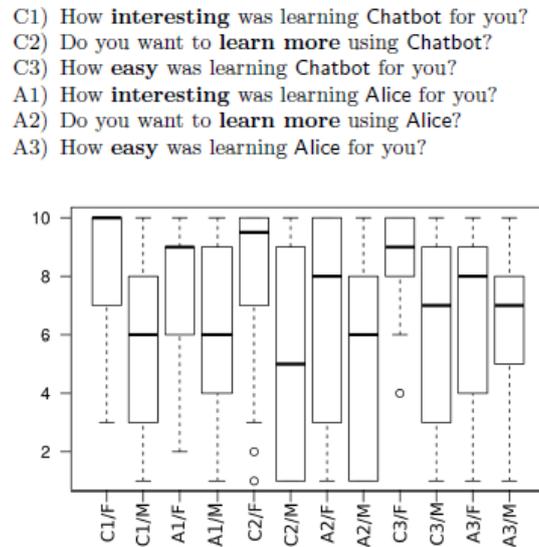


Figura 1 – Resultados dos indicadores obtidos no estudo piloto do ensino secundário organizado por gêneros. Todas as questões são baseadas em uma escala que varia de 0 (significado de “not at all”) a 10 (significado de “very much”).

(Fonte: Benotti et al., 2014)

O diagrama de caixa, apresentado na Figura 1, sumariza os resultados obtidos no inquérito. O interesse pela aprendizagem de ambas as ferramentas foram mais evidentes no sexo feminino. Relativamente à facilidade de aprendizagem, as diferenças observadas não são estatisticamente significativas por género ou por ferramenta, embora seja observada maior variação nas respostas do sexo masculino face ao *chatbot* e no sexo feminino relativamente à ferramenta ALICE.

A Figura 2 mostra o número de vezes que cada palavra foi repetida pelos alunos durante as aulas.

Platform	“Interesting”	“Fun”	“Engaged”
Chatbot	5	6	10
Alice	5	0	8

Figura 2 – Caracterização das opiniões dos agentes de conversação Chatbot e ALICE quanto aos campos Interesting, Fun e Engaged.

(Fonte: Benotti et al., 2014).

2.2. Recuperação de informação Vs. Geração da língua natural

Recentemente o diálogo entre seres humanos e computadores em domínio aberto têm atraído muita atenção no campo do *Natural Language Processing* (NLP), no âmbito dos *chatbots*, por causa das suas importantes aplicações comerciais, e porque evidencia os desafios reais da compreensão e geração da linguagem natural. Ao contrário dos sistemas de diálogo específicos de domínio baseados em regras ou modelos, o diálogo em domínio aberto geralmente requer abordagens baseadas em dados, que podem ser divididas em duas categorias: *retrieval-based dialog systems* e *generation-based dialog systems*, que são sistemas baseados em recuperação da informação e geração da língua natural.

Os *retrieval-based dialog systems* pesquisam o *input*⁸ dado pelo utilizador, denominado por consulta, numa base de conhecimento onde retornam uma resposta adequada. No entanto, não é possível obter novos enunciados, ou seja, todas as respostas precisam de estar na base de conhecimento.

Muitos dos *generation based dialog systems* atuais são baseados em *Recurrent Neural Networks* (RNN) capazes de sintetizar novas respostas, uma RNN captura a semântica da consulta. As redes neurais profundas permitem uma interação complexa por múltiplas transformações não lineares. Apesar disso, as RNNs têm uma grande desvantagem de utilização quando aplicadas aos sistemas de diálogo, onde as frases geradas tendem a ser curtas e sem sentido. Isto ocorre provavelmente porque os *inputs* passados ao *chatbot* são altamente diversificados e pode não conter informação suficiente para uma resposta. Mesmo que tais enunciados possam ser adequados num certo contexto de diálogo, os mesmos podem impulsionar a perda de interesse pela parte do utilizador (Song, Yan, Li, Zhao & Zhang, 2016).

2.3. Frameworks para desenvolvimento de *chatbots*

2.3.1. Microsoft

A Microsoft tem disponível alguns modelos prontos para *bots*⁹ de formulário e *bots* para perguntas e respostas em vários idiomas.

⁸ Fornece dados ao computador.

⁹ Aplicação de *software* desenvolvido para simular ações humanas.

O Azure AD (Azure Active Directory) para além de ser um serviço de gestão de identidade é também uma diretoria *cloud* da Microsoft que pode ser acedida em qualquer lugar. O Azure AD juntamente com o serviço de compreensão de idiomas permite aos programadores desenvolver interfaces conversacionais para vários contextos de utilização (banca, turismo, entretenimento, entre outros...).

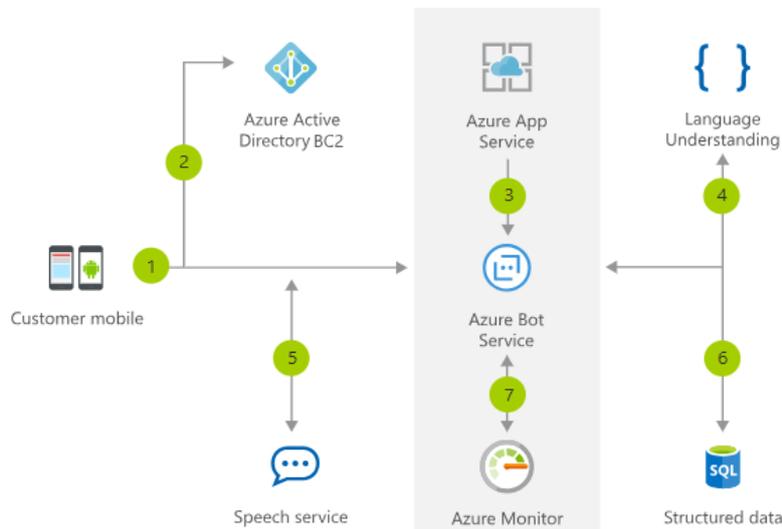


Figura 3 – Fluxograma de funcionamento de um bot da Microsoft.

(Fonte: <https://azure.microsoft.com/pt-pt/services/bot-service/>)

A Figura 3 representa um exemplo de um fluxograma de funcionamento de um *chabot* da Microsoft, onde o utilizador usa a aplicação móvel (etapa 1). Na etapa 2, o Azure AD na vertente do *Business to Customer* (B2C) tem como objetivo a validação do processo de autenticação do programador. A partir do *Application Bot* personalizado (etapa 3), o utilizador pede informações, onde os serviços cognitivos (etapa 4) prestam uma assistência no âmbito do processamento da linguagem natural do pedido. Na etapa 5, a resposta é apresentada ao utilizador e assim se o mesmo estiver satisfeito o *Application Bot* (etapa 6) irá atualizar o pedido do utilizador. Na última etapa o *Application insights* recolhe informações sobre a utilização do desempenho e utilização do *bot*.

O *Language Understanding Intelligent Service* (LUIS) da Microsoft integra-se de forma fluída no serviço de fala e no *Azure Bot Service*, facilitando a criação de um *bot* sofisticado. O LUIS foi projetado para reconhecer informações importantes em diálogos, nomeadamente intenções e entidades.¹⁰

¹⁰ <https://azure.microsoft.com/pt-pt/solutions/architecture/commerce-chatbot/>

2.3.2. Messenger

A Plataforma do Messenger do Facebook é composta por um conjunto de recursos, incluindo *Application Programming Interfaces*¹¹ (APIs), *Plugins* e página de visualização *Web*. Uma dessas APIs é o *bot* do Messenger, que envia mensagens para um utilizador se o mesmo iniciar a conversa em primeiro lugar.

O recurso integrado de NLP do Messenger utiliza a plataforma de linguagem natural wit.ai¹². Tal como na *framework* da Microsoft, o módulo de NLP permite detetar a intenção e o significado das mensagens enviadas aos *bots*. Ao receber uma mensagem, a plataforma devolve o mapeamento das entidades que foram capturadas juntamente com seus dados estruturados, ou seja, o valor de confiança de cada entidade é analisado num valor entre zero e um, indicando a probabilidade do analisador detetar se o reconhecimento está correto.

O NLP está integrado numa caixa de entrada da página do Messenger que possibilita a criação de uma aplicação no wit.ai que por sua vez utiliza o *bootstrap* (*framework front-end open-source* para o desenvolvimento de componentes de interface). Os diálogos entre os utilizadores e o *bot*, são compilados no wit.ai, onde inicialmente contém uma base de dados treinada com elementos linguísticos comuns.¹³

2.3.3. Lex

O Amazon Lex é um serviço da Amazon Web Services para criar interfaces conversacionais de fala e de texto – oferecendo a funcionalidade e flexibilidade do *Natural Language Understanding* (NLU) e do *Automatic Speech Recognition* (ASR),

O NLU tem a tarefa de detetar a quantidade e o tipo de correspondência entre duas frases, sendo que esta é complexa dado que existe muita variabilidade lexical e sintática nas línguas naturais humanas (Fialho, Marques, Martins, Coheur & Quaresma, 2016). O ASR é utilizado em sistemas para converter fala dos utilizadores em texto.

¹¹ Conjunto de instruções e padrões de programação para acesso a uma aplicação de *software*.

¹² Plataforma de linguagem natural *open-source* (<https://wit.ai/docs/quickstart>).

¹³ <https://developers.facebook.com/docs/messenger-platform/built-in-nlp>.

Este serviço permite a qualquer programador criar de forma rápida os seus *chatbots* de conversação, sendo apenas necessário especificar o fluxo de conversação básico na consola do Amazon Lex.

Apesar destas limitações, o Amazon Lex democratiza tecnologias de aprendizagem profunda disponibilizando os recursos do Amazon Alexa, assistente virtual inteligente do Amazon, para todos os programadores.¹⁴

2.3.4. Watson

O IBM Watson representa a combinação das capacidades de NLP, geração e avaliação de hipóteses e aprendizagem dinâmica, tornando-a única para fornecer respostas diretas e baseadas em confiança. O NLP ajuda a compreender a complexidade da informação não estruturada, que compõem até 80 por cento dos dados do mundo de hoje. A geração e avaliação de hipóteses através da aplicação de análises avançadas serve para pesar e avaliar um painel de respostas baseado apenas em evidências relevantes. Por sua vez, a aprendizagem dinâmica ajuda a melhorar a aquisição com base nos resultados para obter um sistema de informação mais inteligente em cada iteração.

O IBM Watson é um sistema de NLP profundo, que procura avaliar o máximo de contexto possível da pergunta a partir da base de conhecimento que está disponível para encontrar respostas.

A navegação eficaz através da atual inundação de informação não estruturada, exige uma nova era de computação a que chamamos sistemas cognitivos e a Watson é um exemplo de um sistema cognitivo, que analisa passagens de texto com precisão e em velocidades e escalas que são muito mais rápidas e muito maiores do que qualquer pessoa pode fazer sozinha.

O Watson não compreende realmente as palavras individuais na língua, mas compreende as características da linguagem que são utilizadas pelas pessoas, e a partir destas determina de uma passagem de texto com um alto nível de precisão.

O Watson separa a questão das possíveis respostas na sua base de conhecimento constituída por pares de interação de perguntas e respostas e, em seguida, examina a

¹⁴ <https://aws.amazon.com/pt/lex/>

questão e o contexto da mesma, utilizando os resultados para estimar um grau de confiança na sua interpretação da pergunta e das potenciais respostas.

Os vários pontos de integração da Watson encontram-se apresentados na Figura 4. Os utilizadores interagem a partir de uma plataforma de sistema de mensagens de redes sociais existentes, como o Messenger da Facebook ou o Slack. A interface do *chatbot* está situada no IBM *cloud*, onde está capacitada para ser uma aplicação móvel ou *chatbot* com uma interface da fala.¹⁵

O *chatbot* recebe a entrada do utilizador e passa-a para a fase da qualificação de diálogo onde interpreta a entrada do utilizador e reúne informações para responder.

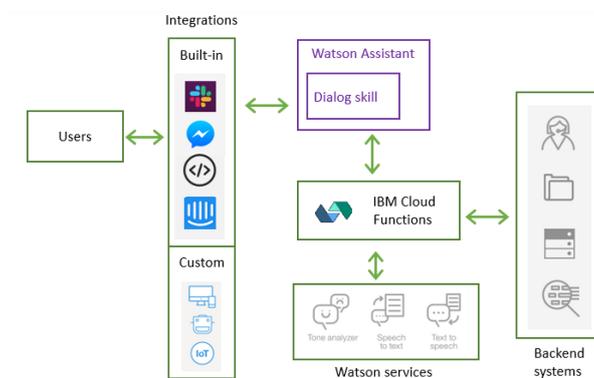


Figura 4 –Diagrama de funcionamento da Watson.

(Fonte: <https://cloud.ibm.com/docs/services/assistant>)

Ao receber a questão, a Watson analisa a maior parte das características da pergunta, gerando um conjunto de hipóteses de uma base de conhecimento que possam ser uma possível resposta. A partir daí, é realizada uma comparação profunda da pergunta e de cada resposta potencial usando centenas de algoritmos de raciocínio, cada um dos quais faz uma comparação diferente, desde correspondência de termos e sinónimos a características temporais e espaciais até fontes relevantes de informação contextual. Cada algoritmo de raciocínio produz uma ou mais pontuações, indicando até que ponto a resposta potencial é apropriada à pergunta. Cada pontuação resultante é então ponderada em relação a um modelo estatístico que é usado para resumir um nível de confiança que o Watson tem sobre a evidência de que a resposta é adequada à pergunta. O Watson repete este processo para cada uma das respostas dos candidatos até que possa encontrar as respostas que mais se aproximam ao resultado esperado pelo utilizador (High, 2012).

¹⁵ <https://cloud.ibm.com/docs/services/assistant>

2.3.5. Diagramflow

Os agentes de conversação são descritos com módulos de NLU, os quais podem ser incluídos numa aplicação, plataforma *online*¹⁶, produto ou serviço e traduzir solicitações de texto. Essa tradução ocorre quando a expressão de um utilizador corresponde a uma intenção no agente de conversação, sendo de seguida entregue uma resposta ao utilizador. Essa resposta pode ser um texto simples ou uma resposta da *webhook*¹⁷. A integração com *webhook* permite passar informações de uma entrada do utilizador para um serviço HTTP (*HyperText Transfer Protocol*) obtendo um resultado a partir de uma plataforma *online* (Vegesna, Jain & Porwal, 2018).

Por exemplo, o Dialogflow envia uma solicitação de *webhook* com os parâmetros de local e data para um serviço meteorológico de terceiros, como se pode visualizar na Figura 5. Este serviço meteorológico retorna uma resposta *webhook* no formato JSON¹⁸. O agente analisa os dados JSON e fornece uma resposta ao utilizador com as informações relevantes.

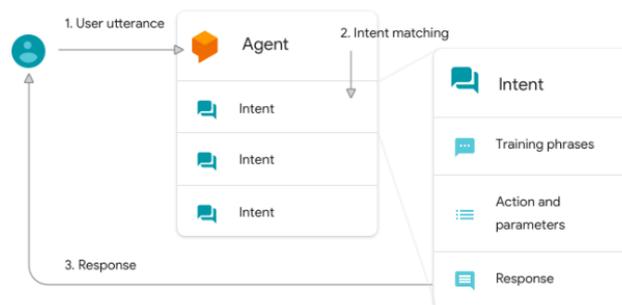


Figura 5 – Exemplo de um Dialogflow's da Google com solicitação *webhook*.

(Fonte: <https://dialogflow.com/docs/intro>)

Os agentes da Dialogflow usam algoritmos de aprendizagem automática para compreender enunciados em linguagem natural, combiná-los com intenções e extrair dados estruturados.

Para além dos modelos de linguagem incorporados no Dialogflow, o agente também aprende com as perguntas do utilizador. Com base nesses dados, cria um algoritmo para

¹⁶ Ligação direta ou remota a um computador.

¹⁷ Retorno de chamada HTTP

¹⁸ Formato leve de troca de informações ou dados entre sistemas.

tomar decisões sobre que intenção deve ser correspondida com a expressão do utilizador.¹⁹

2.3.6. Rasa

O Rasa é um conjunto de ferramentas de aprendizagem automática em *open-source* para os programadores criarem assistentes de inteligência artificial e *chatbots* contextuais. Este conjunto é constituído pelo Rasa NLU e o Rasa Core que fazem a ligação entre a investigação e a aplicação, trazendo avanços recentes no âmbito de aprendizagem automática para programadores menos experientes que pretendem implementar diálogos com inteligência artificial. O Rasa NLU é uma biblioteca para a compreensão de linguagem natural com classificação e extração de entidades e o Rasa Core é uma *framework* de *chatbots* que gere diálogo baseada em aprendizagem automática.

O Rasa é atualmente utilizado por milhares de programadores em todo o mundo. Tal como acontece com muitos outros sistemas de conversação, as ferramentas estão divididas em compreensão da linguagem natural (Rasa NLU) e gestão de diálogo (Rasa Core). O Rasa também está habilitado para migrar dados a partir do Google Diagramflow, Microsoft Luís, Amazon Lex, entre outros.

A arquitetura do Rasa é modular por desenho, permitindo deste modo uma fácil integração com outros sistemas. Por exemplo, o Rasa Core pode ser utilizado como gestor de diálogo em conjunto com serviços NLU que não o Rasa NLU. Enquanto o código é implementado em Python²⁰, ambos os serviços podem expor APIs por HTTP para que possam ser facilmente utilizados por projetos que utilizam outras linguagens de programação.

Quando uma mensagem do utilizador é recebida, o Rasa executa um conjunto de etapas, conforme descrito na Figura 6. A etapa 1 é executada pelo Rasa NLU e todos os passos subsequentes são tratados pelo Rasa Core.²¹

¹⁹ <https://dialogflow.com/docs>.

²⁰ Linguagem de programação de alto nível, usualmente utilizada para processamento de dados.

²¹ <https://rasa.com/docs/>

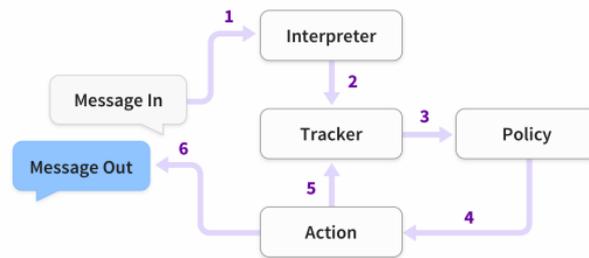


Figura 6 – Diagrama de funcionamento do Rasa.

(Fonte: Bocklisch et al.,2017)

De acordo com a Figura 6, no primeiro passo, o sistema recebe uma mensagem e é passada para um interpretador extrair a intenção, entidades e qualquer outra informação estruturada. Em seguida, na segunda etapa, o *Tracker* mantém o estado da conversa e recebe uma notificação de que uma nova mensagem foi recebida. A *Policy*, no passo 3, recebe o estado atual do *Tracker* e, no passo 4 escolhe que ação que será tomada em seguida. No passo 5, a ação escolhida é registada pelo *Tracker* e no passo 6 a ação é executada incluindo o envio de uma mensagem para o utilizador (Bocklisch, Faulkner, Pawlowski & Nichol, 2017).

2.4. Chatbots na língua portuguesa

2.4.1. O agente conversacional para o Palácio de Monserrate

O Edgar é um *chatbot* desenvolvido pela Laboratório de Sistemas de Língua Falada (L2F) com o objetivo de responder às perguntas colocadas pelos visitantes sobre o Palácio de Monserrate através da Interação com Personagens Virtuais.

A componente NLU do Edgar foi desenvolvida e denominada por *Language Understanding Platform* (LUP). O LUP é gerado por *frames* que são compostas por *slots* preenchidos com um determinado valor. Existem dois algoritmos para o fazer, um deles é o *Frame First*, onde o tipo de *frame* é determinado e os valores de atributo são preenchidos. O segundo, *Attributes First*, analisa os atributos e seleciona o tipo de *frame* que corresponde melhor a estes valores, reforçando a confiança da sua previsão.

Um dos métodos de classificação é gerado em medidas de similaridade, uma técnica que usa a similaridade de palavras como uma medida de classificação. Para fazer isso, o algoritmo passa por exemplos de treino, onde calcula individualmente a similaridade do

input do utilizador. São usadas três medidas de similaridade: *Jaccard*²², *Overlap*²³ e *Dice*²⁴.

Todas as técnicas de classificação disponíveis requerem um *corpus* de treino. Os desempenhos dessas técnicas são sensíveis às características desse *corpus*. Existe um conjunto de ferramentas que realizam o pré-processamento do *corpus* com o objetivo de reduzir a esparsidade. Estas ferramentas incluem a remoção de *Stop Words*²⁵ e *Part-Of-Speech Tagging*²⁶ (POS *Tagging*).

O Edgar possui um outro módulo integrado denominado por Talkpedia, cujo objetivo é gerar respostas inteligentes para as interações *Out of Domain*²⁷ (OOD), usando a *Wikipédia* como base de conhecimento. A avaliação realizada do Talkpedia conta com a participação dos utilizadores através de um questionário depois de terem testado o sistema, em que é dada uma pontuação de 1 a 5 de acordo com a satisfação dessas respostas, sendo 5 a melhor pontuação. Os utilizadores não sabem se a sua experiência se encontra no grupo de controlo e se a sua pergunta foi respondida por um ser humano. Os resultados mostram que as respostas do Talkpedia têm uma pontuação média de 3,00, as de valor “Não sei responder” uma pontuação de 2,20 e as respostas escritas pelo ser humano uma pontuação de 4,20. Este estudo demonstra que os utilizadores preferem que as questões do OOD sejam respondidas mesmo que não estejam completamente corretas, do que as respostas neutras do tipo “não sei” (Fialho *et al.*, 2013).

2.4.2. O agente conversacional baseado em legendas de filmes

O Filipe é um *chatbot* desenvolvido no INESC-ID, cujo objetivo é lidar com OOD usando legendas de filmes. Filipe é baseado no *Say Something Smart* (SSS), um mecanismo de diálogo, onde dada a solicitação do utilizador, o SSS procura para um conjunto de candidatos a resposta em um *corpus* de interações e retorna a melhor resposta

²² Coeficiente que mede a similaridade entre dois conjuntos A e B como $Jac(A,B) = \frac{|A \cap B|}{|A \cup B|}$.

²³ Coeficiente que mede a similaridade a partir da sobreposição de dois conjuntos finitos $overlap(A,B) = \frac{|A \cap B|}{\min(|A|, |B|)}$

²⁴ Coeficiente que compara a similaridade entre dois conjuntos finitos $dice(A,B) = 2 \frac{|A \cap B|}{|A| + |B|}$.

²⁵ Palavra considerada irrelevante que é removida do texto.

²⁶ Processo de marcação de uma palavra em texto (*corpus*) como correspondendo a uma determinada parte do discurso.

²⁷ Respostas pré-definidas para perguntas que são fora do domínio do sistema.

de acordo com uma combinação de vários critérios configuráveis (Mendonça, Melo, Coheur & Sardinha, 2017).

A base de conhecimento é o *Subtle Corpus* para solicitações de OOD, um *corpus* que contém pares de interação e resposta extraídos dos arquivos das legendas do *OpenSubtitles*²⁸. O *chatbot* Filipe usa as legendas de filmes para construir uma base de conhecimento a partir do zero. As principais vantagens deste exemplo são os seus recursos. Amplamente disseminadas de forma gratuita na internet, as legendas são disponibilizadas em várias línguas.

O *Subtle Corpus* é uma coleção de interações extraídas de filmes de diferentes categorias – terror, ficção científica, *western*²⁹ e romance – em Língua Portuguesa.

O SSS recebe a solicitação do utilizador e retorna uma resposta mediante a base de conhecimento do agente de conversação. Este processo consegue comparar uma grande quantidade de expressões, devido à grande quantidade de interações disponíveis. Assim, é feita uma prévia etapa de filtragem no SSS antes de selecionar a resposta mais adequada.

Neste processo, a informação é indexada usando o *Lucene*, uma biblioteca de mecanismo de procura de texto de alto desempenho e *open-source*, amplamente utilizada pelo NLU. Em seguida, aquando solicitação por parte do utilizador, o mecanismo de pesquisa do *Lucene* também é utilizado para devolver o conjunto já classificado de interações. Os resultados devolvidos pelo *Lucene* são baseados num algoritmo interno de pontuação, que leva em consideração tanto as palavras presentes nas interações como no pedido do utilizador.

Para escolher uma resposta do conjunto de interações recolhidas, o SSS executa duas tarefas sequenciais. Primeiro, a filtragem das interações, escolhendo apenas as palavras-chave semelhantes à solicitação do utilizador, conforme os coeficientes de similaridade – os coeficientes de similaridade de *Jaccard* e de *Overlap*. Todas as interações recuperadas acima do limite são mantidas e as restantes descartadas. O coeficiente *Overlap* é definido com uma pontuação mínima de 0,40 (40%) e para o coeficiente de *Jaccard* de 0,70 (70%). Se nenhuma interação for selecionada, então é dada uma resposta definida anteriormente para essa situação, neste caso, “*I’m sorry but I do not know how to answer your question.*”.

²⁸ Plataforma *open-source online* de legendas.

²⁹ Filmes de cowboys

Na segunda tarefa, são analisadas as interações da primeira filtragem. Destas, só uma será selecionada como resposta final. As respostas são comparadas de acordo com a medida de similaridade mencionada anteriormente. É avaliada a concisão de cada resposta. Quantas mais frequências uma resposta tiver, mais essa resposta é plausível para a questão formulada.

Caso nenhuma resposta se diferencie, de acordo com os parâmetros estipulados das restantes respostas, é devolvida uma resposta aleatória – permitindo, por vezes, que este *chatbot* forneça diferentes respostas para o mesmo pedido. Caso contrário, a resposta com maior valor é retornada.

De forma a avaliar o nível de satisfação desta ferramenta por parte dos utilizadores, foi realizado um questionário com base nas respostas fornecidas pelo SSS. Para testar a concordância foi usada a medida de *Cronbach*³⁰ (Maroco & Garcia-Marques, 2006). Obteve-se uma percentagem elevada, tanto para o SSS (0,80), como para o manual (0,84) (Ameixa *et al.*, 2014).

2.5. Chatbots na língua estrangeira

2.5.1. O agente conversacional do Exército

O *Sergeant* (SGT) Blackwell é um *Embodied Conversational Agents* (ECA) que foi projetado numa tela transparente com o objetivo de responder às questões sobre o seu papel no Exército e as novas tecnologias desenvolvidas pelo *Institute for Creative Technology*, nomeadamente o seu criador.

A interação com o agente de conversação é realizada através do diálogo de conversação. A fala do utilizador é convertida em texto a partir de um sistema de ASR e, em seguida, um classificador analisa a saída de texto e seleciona a resposta com a melhor pontuação. O SGT Blackwell tem 55 respostas pré-gravadas para lidar com interações no domínio e 17 para lidar com respostas fora do tópico.

O classificador é baseado em técnicas estatísticas de modelação de linguagem para a recuperação de informações em vários idiomas, e representa uma cadeia de texto com uma distribuição de probabilidade sobre as palavras na cadeia.

³⁰ Mede correlação entre as respostas de um questionário através da análise das respostas dadas pelos avaliadores.

As perguntas e as respostas são vistas como duas linguagens diferentes, a linguagem das questões e a linguagem das respostas. Dada uma pergunta de entrada, o classificador calcula o modelo de linguagem da resposta mais provável para a pergunta usando os dados de treino como um dicionário para “traduzir” a pergunta numa resposta. Em seguida, compara este modelo com o modelo de linguagem das respostas, e seleciona a melhor correspondência.

Em relação ao processo de classificação descrito anteriormente, quando não há correspondência, a questão é considerada como não relacionada a uma das respostas pré-definidas dos 55 conteúdos e o sistema seleciona uma aleatória dentro das 17 respostas disponíveis fora do tópico. Nesta abordagem de escolha aleatória de escolher uma resposta levanta-se a questão de a resposta ser inadequada para a pergunta. Para obter uma melhor correspondência entre a resposta e o tipo de pergunta fora do tópico, as respostas foram agrupadas em classes coerentes. Ou seja, com base em cinco classificadores foram desenvolvidas várias combinações.

Descrição dos classificadores:

- *Baseline* - 55 classes resposta dentro do tópico e uma fora do tópico. As respostas fora do tópico são escolhidas aleatoriamente, então todas estão incluídas na mesma classe;
- *Binary* - Classificador que determina apenas se uma questão está no tópico ou fora do tópico;
- *Off-topic* - O classificador assume que a entrada está fora do tópico e a classifica em uma das classes do OOD descritas acima;
- *On-topic* - O classificador assume que a entrada está dentro do tópico e a classifica em uma das 55 classes relacionadas aos domínios no tópico;
- *Combined* - Este classificador trata as classes que se encontram dentro e fora do tópico como se fossem as mesmas (não há um classificador para distinguir entre o tópico do domínio de entrada / saída, todas as classes estão juntas – classes fora do tópico mais classes no tópico) e classifica a entrada para uma das classes acima referida.

Esta abordagem de lidar com as interações OOD com base em uma taxonomia de tipos de perguntas e respostas fora do tópico é uma maneira de fornecer uma resposta que não responde diretamente à pergunta

As 55 classes foram treinadas com 1000 exemplos, enquanto que a classe fora do tópico foi treinada com 300 exemplos. O teste é composto por 150 questões (100

perguntas dentro do domínio e 50 fora do domínio) que não estão incluídos nos exemplos de treino.

Os resultados foram analisados por três avaliadores utilizando uma escala de 1 a 6, em que 1 corresponde a uma resposta não relacionada com a questão e 6 corresponde a uma resposta perfeita para a questão.

A concordância entre os avaliadores foi avaliada utilizando o *CronBach*. Esta medida apresentou um valor alto, o que demonstra uma alta consistência entre os avaliadores. A avaliação concluiu que a combinação com classificadores *Binary*, *On-topic* e *Off-topic* teve uma média adequada, mostrando que esta técnica deveria ser usada (Leuski *et al.*, 2006).

2.5.2. O agente conversacional Mágico de Oz

Ao contrário dos *chatbots* mencionados anteriormente, o ChatWoz é uma aplicação desenvolvida no INESC-ID em que um utilizador interage com um agente de conversação controlado por um gestor do agente de conversação, sendo que o utilizador não tem conhecimento desse gestor.

O ChatWoz é composto por uma interface controlada pelo gestor, onde é possível controlar vários elementos da interface – como, por exemplo, as emoções do rosto – e onde o gestor estabelece o que o *chatbot* dirá, numa fala sintetizada.

O objetivo deste *chatbot* é recolher dados reais e definir estratégias de diálogo, mas também permite, por exemplo, que crianças com autismo interajam através dos assistentes virtuais de forma a impulsionar as suas capacidades.

O desenho e a implementação do *front-end* do ChatWoz são desenvolvidos em *Unity* – um motor de desenvolvimento utilizado na criação de jogos. Depois do gestor escolher as expressões pré-definidas ou as afirmações de resposta, o texto escolhido será sintetizado.

Na interface principal é apresentado um menu, com as frases gravadas e agrupadas em categorias, que permite acelerar o processo de escolha do que o agente irá dizer. Em vez de escrever, o gestor só tem de clicar no enunciado apropriado, que já foi sintetizado anteriormente.

Em relação às emoções do agente, o gestor pode escolher quatro emoções diferentes em relação à face do agente ao proferir uma sentença. Entre as quais:

- Neutro;
- Alegre;
- Triste;
- Admirado.

Depois de expressar uma dessas emoções, o agente retornará ao seu estado normal (Fialho & Coheur, 2015).

2.5.3. O agente conversacional Max

O agente virtual MAX é um ECA desenvolvido pelo grupo de inteligência artificial da Universidade de Bielefeld cujo objetivo é estudar como se pode modelar o comportamento natural de conversação dos seres humanos e, também, estudar a geração de comportamento multimodal natural.

O Max está equipado com uma câmara que lhe permite captar a passagem dos visitantes e interagir com eles. Os utilizadores interagem com este utilizando um teclado que, apesar de limitativo, evita muitos dos problemas habituais no ASR. O Max interage por fala e por alguns comportamentos não verbais, como expressões faciais ou gestos.

O modelo NLU deste sistema é baseado em atos de diálogo. Cada entrada do utilizador, bem como todas as respostas do agente, são consideradas como um ou mais atos de diálogo.

A interpretação feita pelo Max é dividida em duas fases:

- Na primeira, os conceitos semânticos gerais são identificados (concordância, negação, etc.) por regras de correspondência de padrões;
- Na segunda fase, determina-se a pontuação do *input* do utilizador através das medidas de similaridade.

Quando nenhuma função comunicativa for atribuída, a expressão de entrada é considerada como OOD. O comportamento do agente ao lidar com uma interação OOD é de fornecer uma resposta pré-definida “*I cannot answer your question*” e sugerir um tópico no seu domínio para o utilizador, de forma a manter o diálogo em tópicos em que o Max possa responder.

Por se considerar esta abordagem como monótona, os programadores desenvolveram uma nova abordagem baseada em perguntas e respostas, usando a Wikipédia como base de conhecimento. Uma pontuação é atribuída a cada candidato, e a resposta é selecionada de um conjunto de candidatos classificados em ordem decrescente mediante a pontuação, na qual, o candidato com maior pontuação é considerado como resposta final e correta (Mendes & Coheur., 2013).

Para a avaliação foram utilizadas 200 questões, separando-as por factos, definições e listagens. Nos resultados, o Max respondeu corretamente a 44% das perguntas, pelo que na categoria de definição este respondeu corretamente a 75% das questões mostrando que este sistema alcança bons resultados, quando é abordado por questões desta categoria (Jung & Kopp, 2003).

2.6. Elementos para a avaliação de um *chatbot*

Os *chatbots* sociais de conversação podem ser utilizados para fornecer benefícios às empresas que os utilizam para reduzir o tempo de resposta e fornecer atendimento aprimorado ao cliente de modo a aumentar a satisfação do mesmo. Por outro lado, alguns *chatbots* são especificamente projetados para serem prejudiciais. Tem-se como exemplo as redes de utilizadores falsos, chamados “sybils” no Twitter, que foram implementadas para influenciar artificialmente as contagens de seguidores para aumentar o status social dos utilizadores que as adquirem, espalhar notícias ou rumores falsos e até intimidar utilizadores que expressam determinadas crenças políticas. Na eleição presidencial dos EUA em 2016, até um quinto dos comentários e respostas no Twitter foram impulsionados por contas do Twitter total ou parcialmente autónomas.

Na Figura 7, é possível observar os atributos de qualidade identificados por Radziwill & Benton, (2017) em 32 artigos de conferência e 10 artigos de revista selecionados para o efeito. Em geral as iterações estão alinhadas com o conceito de usabilidade da ISO 9241³¹, onde Abran afirma “A eficácia, eficiência e satisfação com as quais os utilizadores especificados alcançam metas especificadas em ambientes específicos.”. Em particular, a eficácia refere-se à exatidão e integridade com que os utilizadores especificados alcançam suas metas, e a eficiência refere-se a quão bem os

³¹ Norma de usabilidade da Organização Internacional de Normalização que abrange a ergonomia (estudo das organizações de trabalho) da interação entre os seres humanos e os computadores.

recursos são aplicados para alcançar estes objetivos (Abran, Khelifi, Suryan & Seffah, 2003).

EFFICIENCY		
Category	Quality Attribute	Reference
Performance	<ul style="list-style-type: none"> Graceful degradation Robustness to manipulation Robustness to unexpected input Avoid inappropriate utterances and be able to perform damage control Effective function allocation, provides appropriate escalation channels to humans 	<ul style="list-style-type: none"> Cohen & Lane (2016) Thieltges (2016) Kluwer (2011) Morrissey and Kirakowski (2013) Staven (2017)
EFFECTIVENESS		
Category	Quality Attribute	Reference
Functionality	<ul style="list-style-type: none"> Accurate speech synthesis Interprets commands accurately Use appropriate degrees of formality, linguistic register Linguistic accuracy of outputs Execute requested tasks Facilitate transactions and follows up with status reports General ease of use Engage in on-the-fly problem solving Contains breadth of knowledge, is flexible in interpreting it 	<ul style="list-style-type: none"> Kuligowska (2015) Eeuwen (2017) Morrissey & Kirakowski (2013) Wallace (2003) Ramos (2017) Eeuwen (2017) Solomon (2017) Cohen & Lane (2016)
Humanity	<ul style="list-style-type: none"> Passes the Turing test 	<ul style="list-style-type: none"> Weizenbaum (1966); Wallace (2003)
	<ul style="list-style-type: none"> Does not have to pass the Turing Test Transparent to inspection, discloses its chatbot identity Include errors to increase realism Convincing, satisfying, & natural interaction Able to respond to specific questions Able to maintain themed discussion 	<ul style="list-style-type: none"> Ramos (2017) Bostrom & Yudkowski (2014) Coniam (2014) Morrissey & Kirakowski (2013)
SATISFACTION		
Category	Quality Attribute	Reference
Affect	<ul style="list-style-type: none"> Provide greetings, convey personality Give conversational cues Provide emotional information through tone, inflection, and expressivity Exude warmth and authenticity Make tasks more fun and interesting Entertain and/or enable participant to enjoy the interaction Read and respond to moods of human participant 	<ul style="list-style-type: none"> Morrissey & Kirakowski (2013) Pauletto et al. (2013) Solomon (2017) Eeuwen (2017) Ramos (2017) Meira & Canuto (2015)
Ethics & Behavior	<ul style="list-style-type: none"> Respect, inclusion, and preservation of dignity (linked to choice of training set) Ethics and cultural knowledge of users Protect and respect privacy Nondeception Sensitivity to safety and social concerns Trustworthiness (linked to perceived quality) Awareness of trends and social context 	<ul style="list-style-type: none"> Neff & Nagy (2016) Applin & Fischer (2015) Eeuwen (2017) Isaac & Bridewell (2014) Miner et al. (2016) Herzum et al. (2002) Vetter (2002)
Accessibility	<ul style="list-style-type: none"> Responds to social cues or lack thereof Can detect meaning or intent Meets neurodiverse needs such as extra response time and text interface 	<ul style="list-style-type: none"> Morrissey and Kirakowski (2013) Wilson et al. (2017) Radziwill & Benton (2017)

Figura 7 – Atributos de Qualidade de Chatbots e Agentes de Conversação.

(Fonte: Radziwill & Benton,, 2017)

Houve uma consistência notável em todo o material de origem em relação aos atributos de qualidade, embora vários tenham sido repetidos entre as fontes, apenas as fontes

primárias aparecem na Figura 7. No entanto, um atributo estava em questão: se um *chatbot* deve ou não passar no Teste de Turing. De acordo com a maioria dos investigadores, e seguindo as primeiras interfaces de conversação, responder e interagir como um ser humano deve ser a principal prioridade. De facto, este princípio guiou o desenvolvimento por quase quatro décadas desde que o ELIZA ficou *online*.

A posição que um *chatbot* ou agente de conversação não deve ter como objetivo agir como ser humano é compartilhada por alguns analistas que acreditam que o desenvolvimento de caráter e a gestão de caracteres é um dos novos empregos que serão propiciados pela disseminação da inteligência artificial em automação. Como resultado, essa incongruência ilustra o ponto importante de que a lista de atributos de qualidade identificados pela literatura é sugestiva e não prescritiva (Radziwill & Benton., 2017).

2.7. Discussão

Os agentes de conversação podem ser programados para diversos propósitos com base em grandes conjuntos de interações e enfrentam o desafio de escolher uma resposta apropriada face a cada questão.

A construção dos *chatbots* pode ser feita a partir de plataformas *online*, embora neste caso tenham geralmente uma vertente de comércio B2C.

A partir de uma publicação da Gartner³² relativa ao mercado das plataformas conversacionais em 2018, foram retiradas as principais propriedades das frameworks de agentes conversacionais, identificados anteriormente que se encontram representadas na Tabela 1.

Por análise da Tabela 1, verifica-se que a Microsoft é a empresa que oferece uma *framework* mais completa para a construção de um agente de conversação, diferenciando-se também por utilizar o seu próprio módulo de NLP, nomeadamente o LUIS, na plataforma e módulo de NLP de terceiros para uma instalação completa no local.

³² Empresa de pesquisa e consultoria de pesquisa de tecnologia para fornecer recomendações e ferramentas de negócios para as empresas.

Tabela 1 – Frameworks representativos de agentes conversacionais.

(Fonte: <http://msgx.io/pdf/Gartner-reprint.pdf>)

Vendor	Product Name	Deployment			NLP	
		Self-Service Sign up	Cloud	On-Premises ³³	Proprietary	ThirdParty ³⁴
Amazon Web Services (AWS)	Amazon Lex e Alexa for Business	x	x		x	
Facebook	Messenger Platform	x	x		x	
Google	Dialogflow	x	x		x	
IBM	Watson Assistant	x	x		x	
Microsoft	Múltiple ³⁵	x	x	x	x	x
Rasa	Rasa Stack and Rasa Platform	x		x	x	

Quanto ao tipo de *chatbot*, dado que este projeto assenta sobre a base de conhecimento do Ciberdúvidas, o agente de conversação será desenvolvido com base na abordagem *retrieval-based*.

O *PALbot* será utilizado para um contexto educacional, sendo que quanto maior for a procura, maior será a sua base de conhecimento, tendo em conta que a mesma será gradualmente aumentada e atualizada.

Apesar de existirem diversas *frameworks*, como acima referido, este projeto passa por desenvolver um *chatbot* para incorporar na plataforma do Ciberdúvidas, onde é necessário criar um *script* inteiramente customizado e com condutas bem definidas para este propósito. Por esta razão, a construção do *chatbot* por *framework* não é viável.

Comentando agora os casos existentes e começando pelos *chatbots* que estão capacitados para interações OOD, o Edgar tem uma abordagem interessante ao usar o *Talkepedia*, tendo como base de conhecimento a Wikipédia, enquanto o SGT Blackwell tem um número limitado de respostas para responder às perguntas fora do domínio.

³³ Instalado e executado nos computadores do utilizador ou organização.

³⁴ Serviços requisitados de terceiros.

³⁵ Microsoft oferece vários produtos: Bot Framework, BotBuilder Software Development Kits, Language Understanding Intelligent Service (LUIS) e entre outros.

De acordo com a abordagem referida acima do Edgar, os utilizadores preferiam que uma resposta fosse dada mesmo que estivesse errada.

No caso do Max, quando é confrontado com perguntas fora do domínio, o resultado será sempre uma resposta que já foi pré-definida.

O PALbot irá ter o mesmo comportamento.

Quanto ao *corpus*, o Filipe tem uma abordagem diferente dos restantes *chatbots* que já têm respostas definidas. O Filipe constrói o seu *corpus* desde o início, apesar de partir de um conjunto de legendas já definido, ou seja, quanto mais questões forem feitas, mais respostas haverão no seu *corpus*. E como a sua base de conhecimento é aglomerada por legendas de filmes, terá sempre resposta para qualquer pergunta mesmo que esta não vá ao encontro da questão colocada. Os restantes *chatbots* terão sempre que lidar com interações OOD ou com uma fonte *online* ou respostas previamente preparadas para tal.

Atualmente, a recuperação de informações pode ser realizada de forma simples e rápida com o uso de mecanismos de pesquisa permitindo que os utilizadores especifiquem os seus critérios de pesquisa, bem como palavras-chave específicas para obter resultados desejados. Neste sentido, uma medida de similaridade entre palavras-chave e termos de índice é essencial para facilitar prontamente os resultados necessários (Mendonça *et al.*, 2017). A estimação de similaridade usada no Filipe teve um peso mínimo de 40% para o coeficiente de *Overlap* e de 70% para o coeficiente de *Jaccard*. Com estes valores foram obtidos resultados de sucesso, e os seus parâmetros podem ser considerados para utilização futura neste projeto.

O ChatWoz tem também uma abordagem interessante, uma vez que o seu foco passa por definir estratégias para criar diálogos, nomeadamente para ajudar as crianças a interagirem entre si.

O estudo referido acima demonstra a aplicação de um *chatbot* na comunidade do ensino secundário, na qual os alunos ficaram curiosos e interessados no tema. Seria interessante realizar um estudo aplicado ao Ciberdúvidas.

Em suma, os elementos para a avaliação de um *chatbot* demonstram que para desenvolver ou adaptar um *chatbot*, é essencial entender que cada implementação do *chatbot* priorizará diferentes atributos de qualidade em diferentes fases do ciclo de vida do sistema.

Capítulo 3 – PALbot

3.1. Introdução

O Ciberdúvidas é um projeto sem fins lucrativos fundado pelos jornalistas José Mário Costa e João Carreira Bom. Inicialmente, contou com o patrocínio dos Correios de Portugal (CTT) e da Fundação Vodafone de Portugal até Outubro de 2012 e 2013, respetivamente. Desde 2015, este projeto ficou alojado num dos servidores do ISCTE-IUL formando assim uma parceria com esta instituição pública portuguesa de ensino superior, que visa a concretização de projetos comuns focados na língua portuguesa.

O *corpus* é um elemento fundamental para o presente projeto, uma vez que neste se encontra localizada a fonte de conhecimento do Ciberdúvidas a que o PALbot recorre.

Como referido anteriormente no Capítulo 1, neste projeto foi utilizada a base de conhecimento da plataforma Ciberdúvidas que contém 25.785 perguntas e as respetivas respostas para testar diferentes versões do PALbot.

Ao longo deste projeto, foram estudados diferentes tipos de *chatbot*, concluindo que os que melhor se adaptam ao PALbot são o programa AB da ALICE e o SSS do Filipe, adaptando-os à área em estudo.

O programa AB é um interpretador da linguagem *Artificial Intelligence Markup Language* (AIML) que é uma ferramenta de fácil acesso e compreensão. Pode ser criada e editada num editor de texto simples. A sua utilização garante a interatividade com o utilizador, uma vez que com o *Pattern Recognition* é possível personalizar perguntas e respostas com a criação de variáveis, referências de perguntas, entre outros. Deste modo, o PALbot ficaria capacitado a responder perguntas de forma criativa e com personalidade.

O SSS, para além de ter acesso ao algoritmo do Filipe, já foi testado com sucesso, de acordo com o supracitado no capítulo 2. É de salientar que este algoritmo já foi também testado com a língua portuguesa e obteve bons resultados (Ameixa *et al.*, 2014). Uma vez que este projeto passa por abordar a mesma Língua, seria interessante observar o comportamento do SSS com o *corpus* deste projeto.

É fundamental que o PALbot esteja preparado para receber qualquer tipo de perguntas e respostas de forma a aumentar o seu conhecimento.

3.2. Processamento de dados

Os dados do AIML e SSS foram tratados de formas distintas. O primeiro foi sujeito a um processamento mais delicado, onde foi necessário ter um maior cuidado com as palavras da língua portuguesa, nomeadamente sinónimos, antónimos e plurais para que o *input* do utilizador corresponda às marcações de AIML. Já no caso do SSS, apenas foi realizada a transformação do ficheiro Excel³⁶ num ficheiro de formato .txt³⁷ com parâmetros definidos, como se pode observar na Figura 8.

```

1 SubID - 0
2 DialogId - 0
3 Diff - 0
4 IDCiberduvidas - 11238
5 ID - 0
6 I - Substantivos comum de dois, sobrecomum e epiceno
7 R - Um substantivo comum de dois géneros/géneros frequentemente chamado simplesmente comum de dois é
8
9 SubID - 0
10 DialogId - 0
11 Diff - 0
12 IDCiberduvidas - 11238
13 ID - 1
14 I - Eu gostaria de saber exemplos de substantivo comum de dois géneros, sobrecomum e epiceno.
15 R - Um substantivo comum de dois géneros/géneros frequentemente chamado simplesmente comum de dois é

```

Figura 8 – Formato das perguntas e respostas do SSS.

Como referido no capítulo 2, o SSS estava programado para responder com legendas na língua Inglesa. No entanto, tendo em conta o objetivo do desenvolvimento do projeto na língua materna do ISCTE-IUL, as variáveis das linhas 1 a 3 não são importantes neste contexto, pelo que são inicializadas a zero para todas as interações. As variáveis mais importantes, I e R, funcionam como *input* para as medidas de similaridade e output para o utilizador. No Filipe, o I e o R são duas legendas consecutivas, respetivamente, a questão e a resposta. As variáveis IDCiberduvidas e ID são também importantes. O IDCiberduvidas é uma identificação responsável pela equipa da plataforma de Ciberdúvidas. O ID é uma identificação do assunto (ID igual a zero) e da pergunta (ID igual a um) de cada IDCiberduvidas como se pode observar na Figura 8. De forma a existir uma maior correspondência de similaridade independentemente da diversidade da pergunta do utilizador, os IDs 0 e 1 têm a mesma resposta com perguntas diferentes, mas da mesma natureza.

De modo a obter as perguntas e respostas no formato correto para permitir a sua leitura com o SSS, estes devem passar por duas fases.

³⁶ Editor de folhas de cálculo produzido pela Microsoft

³⁷ Arquivo de texto

Inicialmente o *corpus* foi sujeito a um processo de limpeza a partir do *script* desenvolvido em Python denominado por *CleaningForSSS*, ilustrado Figura 9. Este *script* usa o *beautiful soup*³⁸ para retirar as marcações de HTML e outros caracteres indesejados. Em seguida desenvolveu-se um script, também em Python, para gerar perguntas e respostas no formato ilustrado na Figura 10.

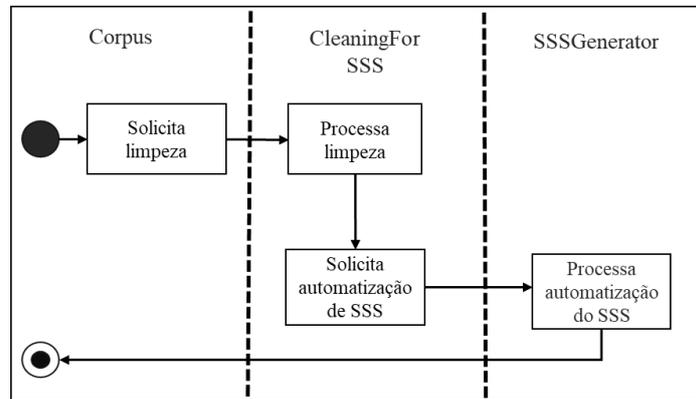


Figura 9 – Diagrama de atividades do PALbot, versão SSS.

	A	B	C	D
1	IDCiberduvidas	Assunto	Pergunta	Resposta
2	11238	Substantivos comum de do	Eu gostaria de saber e	Um substantivo comum de dois géneros/
3	7411	À ou há	À ou há. Gostaria de À	é a contracção da preposição a com o
4	17360	A definição de nódulo sólido	Gostava que me explic	A expressão nódulo sólido hipocogénico
5	27656	Vírgulas com contudo, poré	Segundo me recorde,	Segundo Celso Cunha e Lindley Cintra N
6	10638	Como fazer uma análise cri	Preciso fazer uma aná	A resposta a esta pergunta pode ser ext
7	30126	Co-orientador, ou coorienta	Qual a maneira correc	Segundo o novo Acordo Ortográfico, usa-
8	33650	As abreviaturas de senhor/	Ex.mo/a s.Exmo/a. s.	Em relação às abreviaturas, não parece t
9	14050	20 palavras semanticament	Uma dúvida de alguns	1. Acostamento no Brasil – berma em Pt

Figura 10 – Dados de ciberdúvidas em Excel.

Relativamente ao AIML, foi também necessário desenvolver um *script* com um objetivo idêntico ao processo do SSS, denominado por *CleaningForAIML*, ilustrado na Figura 11. De seguida, foi gerado um ficheiro em formato AIML já com os pares perguntas e respostas preparadas para ser inserido no programa AB e simultaneamente foi gerado um ficheiro em formato txt com palavras sinónimas e plurais em relação às frases presentes nas marcações do *input* do utilizador, uma vez que o programa AB não se encontra adaptado à Língua Portuguesa.

³⁸ Biblioteca de Python para analisar HTML e XML.

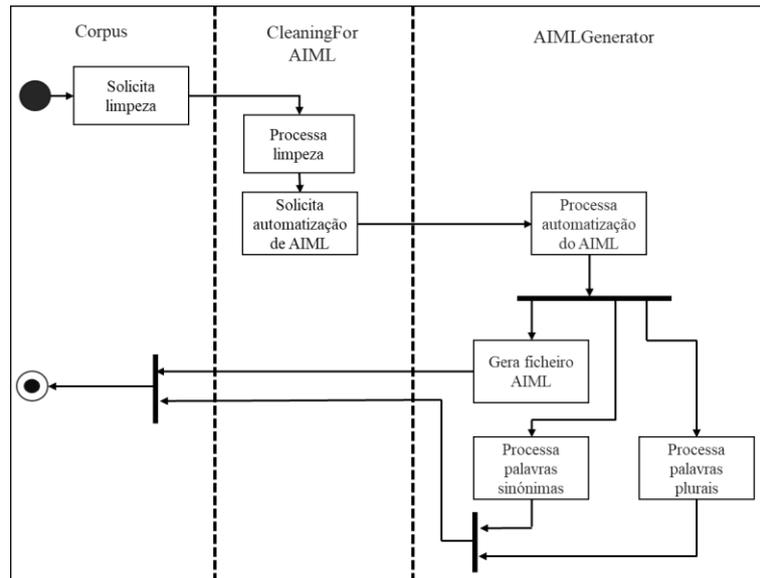


Figura 11 – Diagrama de atividades do PALbot, versão AIML.

De forma a haver correspondência, foi ainda necessário criar estrategicamente várias entradas para a mesma frase no ficheiro gerado em AIML, como se pode visualizar no padrão ilustrado na Figura 12.

```

1  <?xml version="1.0" encoding="UTF-8" ?>
2  <aiml>
3
4  <category>
5    <pattern>À OU HÁ</pattern>
6    <template>À é a contracção da preposição "a" com o artigo definido no feminino singular "a". Ex.1. Vou à escol
7  </category>
8  <category>
9    <pattern>* À OU HÁ</pattern>
10   <template><srail>À OU HÁ</srail></template>
11 </category>
12 <category>
13   <pattern>À OU HÁ *</pattern>
14   <template><srail>À OU HÁ</srail></template>
15 </category>
16 <category>
17   <pattern>* À OU HÁ *</pattern>
18   <template><srail>À OU HÁ</srail></template>
19 </category>
20
21 </aiml>

```

Figura 12 – Exemplo de correspondência de padrões no AIML.

3.3. Versão baseada na ALICE

3.3.1. Visão Geral

O AIML permite que os programadores insiram conhecimento em *chatbots* baseados na tecnologia de *software* livre ALICE (Wallace, 2003).

O AIML é baseado em unidades básicas de diálogo formadas por padrões de entrada do utilizador e por respostas do *chatbot*. Essas unidades básicas são chamadas de categorias, e o conjunto de todas as categorias (*category*) compõem a base de conhecimento do *chatbot*. Entre os objetos AIML, a marcação de *category* define uma

unidade de conhecimento, a marcação de *pattern* define uma possível entrada do utilizador e a marcação de modelo define a resposta do *chatbot* para uma determinada entrada do utilizador (Marrieto, 2013).

Ao longo dos anos foi surgindo uma variedade de interpretadores AIML que utilizam Java, Ruby ou Python como linguagem de programação, entre outras. As três melhores opções para criar um *chatbot* são os programas Z, O e AB (McNeal *et al.*, 2013).

O programa Z é a linguagem proprietária da Pandorabots, um dos mais antigos e maiores serviços de *host* de *chatbot* do mundo. Em fevereiro de 2012, o serviço gratuito da Pandorabots guardava de mais de 201.000 *chatbots*.

O programa O é baseado numa aplicação desenvolvida em PHP que utiliza uma base de dados MySQL. As versões mais antigas deste programa careciam de algumas das funcionalidades do programa Z, mas essa lacuna foi reduzida de forma significativa com o lançamento da versão do AIML 2.0. O programa O apresenta as mesmas vantagens e desvantagens do programa Z, nomeadamente a necessidade de configurar o mecanismo no próprio servidor, bem como a necessidade de ter instalado e executado o PHP e o MySQL. Ao utilizar o programa O é ainda necessário ter um domínio próprio na *World Wide Web*³⁹, tornando-se mais complicado de utilizar face ao Pandorabots (McNeal *et al.*, 2013).

O programa AB é um compilador que suporta o AIML 2.0 e é escrito em Java. O AIML 2.0 foi lançado em Janeiro de 2013 pela ALICE AI Foundation e representa uma tentativa de resolver algumas das limitações de versões AIML, mantendo a simplicidade da linguagem (McNeal & Newyear, 2013).

O *PALbot* será desenvolvido com base no programa AB, com o objetivo de aplicar perguntas e respostas no formato AIML aplicadas em âmbito institucional, nomeadamente na plataforma *online* Ciberdúvidas. Em geral, o programa AB contém o algoritmo *Graphmaster* que funciona sobre um conjunto de nós chamados *Nodemappers* e mapeia os ramos de cada nó, onde os ramos são palavras únicas ou *wildcards*. (Wallace, 2003). O número de nós é igual ao número de categorias e cada nó contém os elementos *pattern* e *template*. O *pattern* corresponde ao padrão de busca e o *template* corresponde à resposta (dos Santos *et al.*, 2016).

³⁹ Sistema de documentos que são interligados executados na internet

A Figura 13 representa o diagrama de classes da ALICE que contém a implementação dos métodos que permitem a leitura e análise dos ficheiros AIML e a configuração da diretoria das variáveis.



Figura 13 – Diagrama de classes da ALICE.

3.3.2. Marcações Artificial Intelligence Markup Language

Uma base de conhecimento em AIML é formada por um conjunto de categorias. As categorias são organizadas por temas e armazenadas em arquivos com extensão .aiml, com o objetivo de organizar e facilitar o processo de manutenção da base de conhecimento (Marrieto *et al.*, 2013).

A linguagem padrão AIML é simples, consistindo apenas em palavras, espaços e símbolos *wildcard* – nomeadamente caracteres especiais. As palavras são separadas por um único espaço e os símbolos *wildcard* usam-se para palavras não previstas. As primeiras versões do AIML permitiam apenas um carácter *wildcard* por marcação.

A Figura 14 mostra que cada arquivo AIML começa com a marcação <aiml> e é fechado por uma marcação </aiml>, na linha 2 e 7. Esta marcação contém a versão e os atributos de codificação, conforme definido na linha 1. O atributo *version* identifica a versão AIML utilizada na base de conhecimento. Se este atributo for omitido, não resultará em erros de execução, mas poderá causar confusão durante tarefas de manutenção ou atualização do sistema. O atributo de codificação identifica o tipo de codificação de caracteres que será usada no documento. Neste projeto, para usar o AIML é necessário recorrer à codificação *UTF-8* uma vez que o *corpus* é em Língua Portuguesa.

Dentro do âmbito `<aiml> </aiml>`, linhas 2 a 7, pelo menos um elemento `<category>` deve existir.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml version="2.0">
3  <category>
4      <pattern>Olá ISCTE</pattern>
5      <template>Saudações académicas!</template>
6  </category>
7  </aiml>

```

Figura 14 – Marcação AIML para saudação.

A Figura 14 demonstra um simples código para a marcação `<category>` que corresponde ao diálogo da Figura 15.

```

Utilizador: Olá ISCTE
ISCTE-IUL: Saudações académicas!

```

Figura 15 – Diálogo entre utilizador e ISCTE-IUL.

A marcação `<pattern>` contém uma possível entrada do utilizador. Há apenas um único elemento `<pattern>` em cada marcação `<category>` e deve ser o primeiro elemento a ser definido. Além disso, em cada elemento, as palavras são separadas por espaços únicos e *wildcards* podendo substituir partes de uma frase. Na Figura 14, o código AIML `<pattern> Olá ISCTE </pattern>`, definido na linha 4, indica que o programador do *chatbot* entendeu que a frase “Olá ISCTE” é uma possível entrada do utilizador.

A marcação `<template>`, que deve estar inserida numa marcação `<pattern>` e após uma marcação `<category>`, contém possíveis respostas que o *chatbot* pode dar ao utilizador. A maioria das informações do *chatbot* é limitada por este elemento. Esta marcação pode salvar dados, ativar outros programas, dar respostas condicionais ou até referenciar outras categorias de respostas.

O padrão AIML 2.0 permite vários *wildcards* em cada padrão, mas a linguagem é projetada para ser o mais simples possível.

Os dois tipos de contexto opcionais são chamados *that* e *topic*. A palavra-chave *that* aparece dentro da categoria e o seu padrão deve coincidir com a última expressão do *chatbot*. A palavra-chave *topic* aparece fora da categoria e reúne um grupo de categorias.

A palavra-chave *that* no AIML refere-se à expressão anterior do *chatbot*. Especificamente, se o *chatbot* responder com um parágrafo de múltiplas frases, o valor é definido para a última frase da sequência.

Uma aplicação comum do palavra-chave *that* é normalmente utilizado em perguntas de respostas sim ou não.

De acordo com a Figura 16, esta categoria é ativada quando o utilizador introduz “SIM” como *input* para o *chatbot*. Se o *chatbot* perguntasse “Gostas de filmes?”, essa categoria iria corresponder com a resposta e seria “Qual é o seu filme favorito?”, continuando assim a conversa na mesma linha.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml>
3
4  <category>
5      <pattern>SIM</pattern>
6      <that>GOSTAS DE FILMES</that>
7      <template>QUAL É O TEU FILME FAVORITO?</template>
8  </category>
9
10 </aiml>

```

Figura 16 – Marcação AIML para interação de resposta.

A primeira parte da categoria a ser correspondida é a entrada. Se mais de uma categoria tiver o mesmo padrão de entrada, o programa poderá distinguir entre elas, dependendo do valor de `<that>`. Se duas ou mais categorias tiverem o mesmo `<pattern>` e `<that>`, a seleção final escolherá a resposta com base no `<topic>`, de acordo com o nome do *topic* como se pode ver na Figura 17.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml>
3
4  <topic name="CARROS">
5
6  <category>
7      <pattern>*</pattern>
8      <template>
9          <random>
10         <li>QUAL É O TEU CARRO FAVORITO?</li>
11         <li>QUAL É O CARRO QUE CONDUZES?</li>
12         <li>TENS MUITAS MULTAS?</li>
13         <li>O MEU CARRO FAVORITO É AQUELE EM QUE TENHO CONDUTOR.</li>
14         </random>
15     </template>
16 </category>
17
18 </aiml>

```

Figura 17 – Marcação AIML para respostas aleatórias mediante o *topic*.

Esta estrutura sugere que nunca se utilize `<that>` a menos que existam duas categorias com o mesmo `<pattern>` e que nunca use `<topic>`, a menos que haja duas categorias com o mesmo `<pattern>` e `<that>`. Ainda assim, uma das aplicações mais úteis para `<topic>` é criar linhas de captura dependentes do assunto como se pode ver na Figura 17.

Uma das propriedades mais úteis da linguagem AIML é a possibilidade de direcionar diferentes marcações <pattern>, entradas do utilizador, para uma única marcação <template>. Assim, o interpretador AIML pode procurar eficientemente uma resposta para diferentes tipos de entrada do utilizador. Essa possibilidade é obtida utilizando a marcação <srail>. Essa marcação poderá ter aplicações como a redução simbólica, dividir e conquistar, resolução de sinónimos e deteção de palavras-chave.

3.3.3. Redução simbólica

Na Figura 18, existe uma categoria preparada para falar sobre a localização do ISCTE-IUL (linhas 4-7) e outra categoria preparada para falar sobre a localização do IST (linhas 9-12). O diagrama na Figura 19 ilustra os comandos <pattern> dessas categorias em que a etapa A se refere à localização do ISCTE-IUL e a etapa B se refere ao IST.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml version="2.0">
3
4  <category>
5      <pattern> LOCALIZAÇÃO ISCTE-IUL </pattern>
6      <template>Av. das Forças Armadas, 1649-026 Lisboa</template>
7  </category>
8
9  <category>
10     <pattern> LOCALIZAÇÃO IST </pattern>
11     <template>Av. Rovisco Pais 1, 1049-001 Lisboa</template>
12 </category>
13
14 <category>
15     <pattern> ONDE FICA * </pattern>
16     <template><srail> LOCALIZAÇÃO <star/></srail></template>
17 </category>
18
19 </aiml>

```

Figura 18 – Múltiplas marcações AIML.

Considerando que o utilizador pode perguntar sobre as mesmas entidades de várias maneiras, a aplicação da técnica de redução simbólica inicia-se com a criação de uma nova categoria que modela uma forma variada de perguntar sobre as entidades. Como exemplo tem-se a categoria que está codificada nas linhas 14 a 17 da Figura 18 e considera-se que o utilizador pergunta “Onde fica *”, onde o *wildcard* “*” identifica a entidade que o utilizador deseja saber.

Se o utilizador inserir um texto que corresponde ao padrão definido na linha 15 da Figura 18 e passo C na Figura 19, o comando <srail> LOCALIZAÇÃO <star /> </ srail>

(linha 16) redireciona a resposta do *chatbot* para outra categoria. O comando `<star />` insere o texto capturado pelo *wildcard* na linha 15.

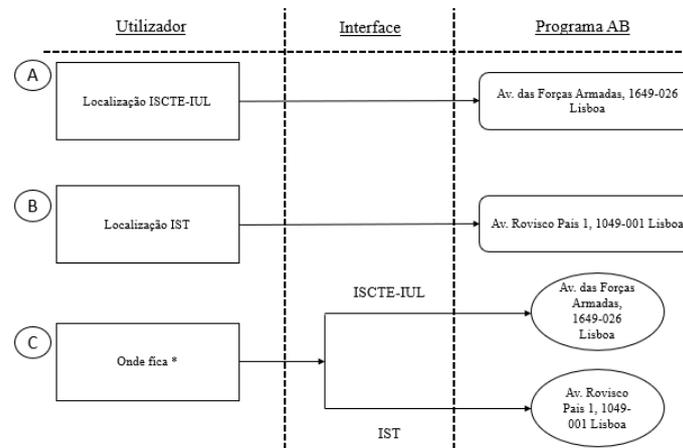


Figura 19 – Diagrama de funcionamento da marcação AIML

(Adaptado de Marietto et al., 2013).

A Figura 20 apresenta um exemplo complementar onde a primeira categoria deteta a palavra-chave quando aparece sozinha e fornece a correspondente resposta genérica. A segunda categoria deteta a palavra-chave como o sufixo de uma sentença e a terceira deteta como o prefixo de uma sentença de entrada, sendo que ambas utilizam `<srail>` para vincular à primeira categoria. Deste modo os três casos produzem sempre uma resposta.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml version="2.0">
3
4
5  <!-- Saudações -->
6  <category>
7    <pattern>OLÁ</pattern>
8    <template>
9      <random>
10     <li>Saudações académicas! </li>
11     <li>Olá </li>
12     <li>Olá em que posso ajudar? </li>
13     <li>Olá em que posso ser útil?</li>
14     </random>
15   </template>
16 </category>
17
18 <category>
19 <pattern>OLA</pattern>
20 <template><srail>OLÁ</srail></template>
21 </category>
22
23 <category>
24 <pattern>OLA * </pattern>
25 <template><srail>OLÁ</srail></template>
26 </category>

```

Figura 20 – Múltiplas marcações AIML para saudação.

3.3.4. Dividir e conquistar

De acordo com Marrieto *et al.* (2013), muitas frases individuais podem ser aglomeradas a duas ou mais frases, assim como a resposta individual pode ser obtida pela combinação das novas respostas. Na linha 10 da Figura 21, a sentença começa com “BOM DIA *”, onde o *wildcard* “*” pode ser qualquer sequência de caracteres. A técnica de dividir e conquistar consiste, neste caso, em que uma entrada do utilizador que comece com a palavra “BOM DIA”, seguida por qualquer *string*, seja tratada pelo comando `<pattern> BOM DIA </pattern>` na linha 5.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml version="2.0">
3
4  <category>
5    <pattern> BOM DIA </pattern>
6    <template> Bom dia, em que posso ser útil? </template>
7  </category>
8
9  <category>
10   <pattern> BOM DIA * </pattern>
11   <template> <srai> BOM DIA </srai> </template>
12 </category>
13
14 </aiml>

```

Figura 21 – Marcação AIML para dois tipos de entrada.

As Figuras 21 e 22 apresentam o código AIML de uma conversa que utiliza esta técnica. A linha 10 da Figura 21 e o passo A da Figura 22 definem um padrão de entrada do utilizador que comece por dizer “Bom dia”, seguida pelo *wildcard*. Neste caso, na Linha 11 da Figura 21 e passo B do fluxograma da Figura 22, o *chatbot* redireciona o sistema de fluxo de execução para a linha 5 com o comando `<srai> BOM DIA </srai>`. A linha 5 e passo C definem o padrão “BOM DIA”. Assim, a resposta na Linha 6 e passo D será a resposta apresentada ao utilizador.

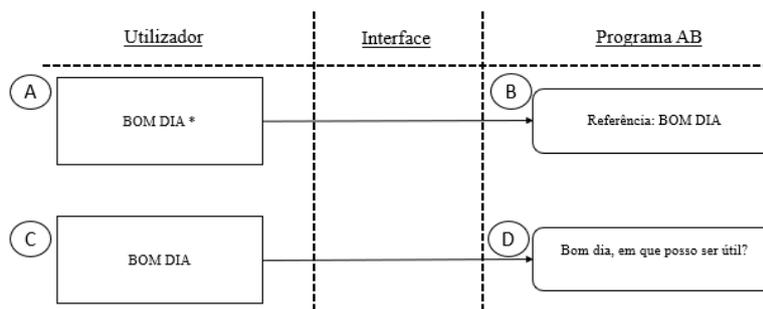


Figura 22 – Diagrama de funcionamento da marcação AIML

(Adaptado de Marrieto *et al.*, 2013).

3.3.5. Redução de sinónimos

O tratamento de sinónimos é muito importante uma vez que podem aparecer diferentes palavras, mas com significados semelhantes, dependendo do contexto. Utilizando a marcação <srai>, o *chatbot* apresenta a mesma resposta para sinónimos, mapeando um conjunto de entradas relacionadas a uma entrada padrão raíz.

A Figura 23 ilustra o uso da técnica de resolução de sinónimos, onde os termos “OLÁ”, “BOM DIA”, “BOA TARDE” e “BOA NOITE” são tratadas como sinónimos. Se o utilizador inserir “Bom dia”, o *chatbot* apresentará a resposta definida nas linhas 10 a 13. A marcação <random> é uma marcação utilizada com o objetivo de responder a uma entrada do utilizador de diferentes maneiras. Cada resposta possível deve ser delimitada pela marcação . Desta forma, as respostas do *chatbot* são tratadas como uma lista, e as respostas são seleccionadas aleatoriamente pela ALICE.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml>
3
4
5  <!-- Saudações -->
6  <category>
7    <pattern>OLÁ</pattern>
8    <template>
9      <random>
10     <li>Saudações académicas! </li>
11     <li> Olá </li>
12     <li> Olá em que posso ajudar? </li>
13     <li> Olá em que posso ser útil?</li>
14   </random>
15 </template>
16 </category>
17
18 <category>
19   <pattern>BOM DIA</pattern>
20   <template><srai>OLÁ</srai></template>
21 </category>
22
23 <category>
24   <pattern>BOA TARDE </pattern>
25   <template><srai>OLÁ</srai></template>
26 </category>
27
28 <category>
29   <pattern>BOA NOITE</pattern>
30   <template><srai>OLÁ</srai></template>
31 </category>
32
33 </aiml>

```

Figura 23 – Múltiplas entradas AIML de saudação para respostas aleatórias.

3.3.6. Set e Get

As marcações <set> e <get> permitem que o *chatbot* trabalhe com variáveis. Existem variáveis predefinidas no idioma para armazenar informações, como dados sobre o *chatbot*. No entanto, as variáveis também podem ser criadas, definindo um nome e valores de inicialização (Marietto *et al.*, 2013).

A marcação `<set>` é utilizada para criar variáveis e deve estar dentro do corpo da marcação `<template>`.

Na Figura 24 observa-se a sua forma generalizada, onde `variableName` é o nome da variável a ser criada ou alterada, e o `variableValue` é o valor a ser atribuído à variável.

```
<set name = "variableName"> variableValue </set>
```

Figura 24 – Definição de variável em AIML.

A Figura 25 ilustra o uso da marcação `<set>`. Neste caso, o nome informado pelo utilizador na frase “O meu nome é *” (linha 9), é identificado pelo *wildcard* “*”. Depois disso, na linha 10, o comando `<set name = “name”> <star /> </ set>` armazena na variável `name`, o texto capturado pelo comando `<star />`, que neste caso é o nome do utilizador.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml version="2.0">
3
4  <category>
5    <pattern>Qual é o meu nome</pattern>
6    <template><get name="name" /></template>
7  </category>
8  <category>
9    <pattern>O meu nome é *</pattern>
10   <template>Prazer em conhecer-te, <set name="name"><star /></set></template>
11 </category>
12
13 </aiml>
```

Figura 25 – Marcação AIML usando variável.

3.3.7. Outros caracteres *wildcard*

Existem dois outros *wildcards*, “_” e “#”. Estes *wildcards* assumem a prioridade mais alta. Mesmo que a entrada forme uma correspondência exata com um padrão, a correspondência poderá ser substituída por um padrão que contenha um destes caracteres.

Se a entrada for “Olá ISCTE!”, de acordo com a Figura 26, o primeiro padrão formará uma correspondência devolvendo a resposta formulada na linha 6. Ou seja, “_” terá sempre prioridade, mesmo que haja uma correspondência exata com um padrão, neste caso na linha 10. Este carácter tem a mesma funcionalidade que o *wildcard* “*” (contém um ou mais caracteres), mas é mais prioritário.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml>
3
4  <category>
5    <pattern>OLÁ _</pattern>
6    <template>Saudações Académicas!</template>
7  </category>
8
9  <category>
10   <pattern>OLÁ ISCTE!</pattern>
11   <template>OLÁ!</template>
12 </category>
13
14 </aiml>
```

Figura 26 – Marcação AIML com *wildcard* _.

Um outro *wildcard* de prioridade é ilustrado na linha 5 da Figura 27 e é representado como “#”. Este *wildcard* significa que a frase pode conter zero ou mais caracteres na posição definida. Ou seja, se o *input* do utilizador for “Olá ISCTE”, o algoritmo irá devolver a resposta de “Olá, como te chamas?” que está definida na linha 6, uma vez que o *wildcard* “#” tem também prioridade em relação à exata correspondência com um padrão.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml version="2.0">
3
4  <category>
5      <pattern>Olá #</pattern>
6      <template>Olá, como te chamas?</template>
7  </category>
8
9  <category>
10     <pattern>Olá ISCTE</pattern>
11     <template>Olá, tudo bem?</template>
12 </category>

```

Figura 27 – Marcação AIML com *wildcard* #.

De acordo com a plataforma Pandorabots a Figura 28 demonstra a prioridade que os *wildcards* têm juntamente com os padrões que contém uma correspondência exata.

Olá # > Olá _ > Olá ISCTE > Olá ^ > Olá *

Figura 28 – Prioridade de *wildcards*.

Além dos *wildcards* “_” e “#”, é possível usar um outro carácter que dará uma correspondência exata e terá prioridade ainda mais alta, substituindo-os assim. A marcação “\$” significa que um padrão será correspondido primeiro com uma palavra específica, como se pode observar na Figura 29.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <aiml>
3
4  <category>
5      <pattern>$Que dia é hoje</pattern>
6      <template>Hoje é <date format="%B %d, %Y" /></template>
7  </category>
8
9  <category>
10     <pattern>_ dia é hoje</pattern>
11     <template>Hoje é <date format="%B %d, %Y" /></template>
12 </category>
13
14 </aiml>

```

Figura 29 – Diferença de *wildcard* \$ e _ na marcação AIML.

O primeiro padrão corresponde a “Que dia é hoje?”, enquanto o segundo padrão corresponde a todas as outras entradas que terminam com “dia é hoje”.

Contudo, o símbolo “\$” não é um *wildcard*. É um marcador que indica que para a palavra especificada, se deve substituir a categoria que teria sido correspondida.

3.3.8. Interações fora do domínio

Neste projeto, os casos que se encontram fora do domínio serão tratados com respostas já preparadas caso o utilizador faça perguntas às quais o agente de conversação não saiba responder, a Figura 30 ilustra como estas respostas se definem em AIML. Para tal, bastará definir o *wildcard* “*” no *pattern* para o algoritmo entenda com que respostas deve lidar caso a pergunta seja OOD. Neste caso as respostas são dadas aleatoriamente como definido nas linhas 8 a 10.

```

1  <<?xml version="1.0" encoding="UTF-8"?>
2  <aiml version="2.0">
3
4  <category>
5      <pattern>*/</pattern>
6      <template>
7          <random>
8              <li>Peço desculpa, mas não percebi o que disse.</li>
9              <li>Não entendi.</li>
10             <li>Não percebi.</li>
11          </random>
12      </template>
13  </category>
14
15 </aiml>

```

Figura 30 – Respostas default na marcação AIML.

3.4. Versão baseada no Say Something Smart

Uma das principais razões para a escolha do SSS é o acesso ao seu algoritmo. Para além do algoritmo, é também importante configurar algumas das suas propriedades e ambiente de execução tais como a linguagem, o ficheiro configurações e a sua base de conhecimento.

A Figura 31 apresenta o diagrama de classes do SSS que contém a implementação dos métodos de identificação das perguntas e respostas, onde irá ser utilizado na atualização do ficheiro em Excel.

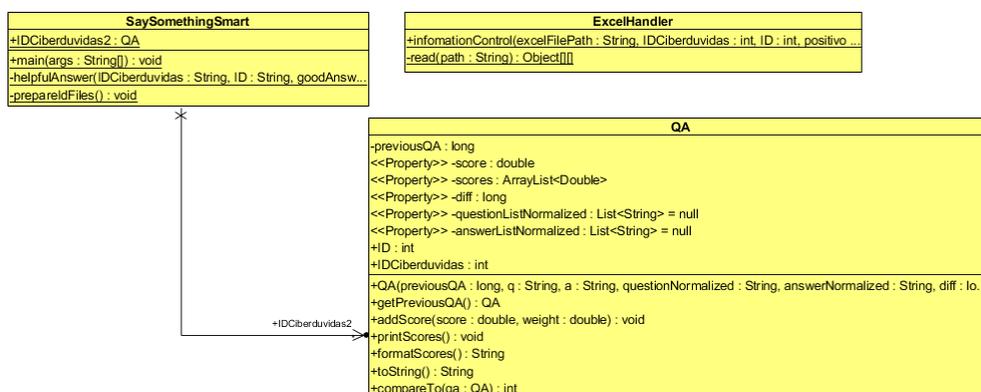


Figura 31 – Diagrama de classes do SSS.

3.4.1. Funcionamento do *Say Something Smart*

Para cada interação recebida no sistema de diálogo SSS, o Lucene irá compará-la com as perguntas que estão no corpus, atribuindo diferentes relevâncias a cada uma delas, e criando uma lista organizada com os pares de perguntas e respostas mais compatíveis. Para filtrar essa saída dada pelo Lucene, cuja dimensão pode ser configurada pelo programador, o SSS pode ser configurado para utilizar como estimador de similaridade uma combinação de diferentes métricas, como as medidas de similaridade de *Jaccard* e de *Overlap*, que calculará as semelhanças entre a entrada do utilizador e os pares de perguntas e respostas retornados pelo Lucene.

O SSS calculará as melhores combinações do Lucene, tentando descobrir qual a melhor resposta com seu próprio algoritmo. Para isso, o SSS utiliza cinco diferentes medidas com diferentes pesos apresentadas na Figura 32, que combinadas devem dar a melhor resposta possível. Cada medida tem um peso diferente e pode ser combinada para fornecer uma lista organizada de cinco respostas com os melhores valores possíveis combinados. Esses pesos podem ser configurados no ficheiro de configuração. Não é obrigatório usar as quatro medidas ao mesmo tempo, mas pelo menos uma deve estar ativa. Cada critério será utilizado para selecionar a melhor resposta, identificado pelo seu nome. Cada critério tem como peso um valor inteiro, entre 0 e 100. A soma dos pesos de critério deve ser igual a 100, como se pode observar na Figura 32 onde após vários testes se chegou ao balanceamento adequado a cada um dos 3 critérios utilizados na seleção de respostas que o *chatbot* dá mediante as perguntas (dos Santos, 2015).

```

34      <!-- Answer selection criteria -->
35      <criteria>
36          <criterion name="AnswerFrequency"
37                  weight="20"/>
38          <criterion name="AnswerSimilarityToUserQuestion"
39                  weight="20"/>
40          <criterion name="QuestionSimilarityToUserQuestion"
41                  weight="60"/>
42          <criterion name="SimpleTimeDifference"
43                  weight="0"/>
44          <criterion name="SimpleConversationContext"
45                  weight="0"
46                  nPreviousInteractions="2"/>
47      </criteria>

```

Figura 32 – Critério na seleção de resposta.

A resposta com maior frequência é o primeiro critério. Esse critério dará um maior valor à resposta mais comum entre as respostas possíveis, de modo a fornecer um valor de redundância ao *corpus*. Ao atribuir a pontuação mais alta à resposta mais frequente, o

SSS assume que se uma resposta é mais frequente para uma determinada questão, então tem uma maior probabilidade de ser a resposta correta.

O segundo critério mede a similaridade entre a resposta e a entrada do utilizador. Uma das muitas maneiras possíveis de enfrentar o problema de dar uma resposta viável é considerar que, se uma resposta tem muitas semelhanças com a questão, é altamente provável que seja uma boa resposta.

O terceiro critério mede a similaridade da pergunta com a entrada do utilizador. Se for possível supor que uma resposta tem algumas semelhanças com uma questão, também é correto pensar que, se houver uma pergunta, na base de conhecimento, muito semelhante à interação do utilizador, podemos supor que qualquer resposta dada a essa pergunta também poderá ser uma resposta possível para a entrada dada pelo utilizador.

O quarto critério considerará a diferença de tempo entre o *input* do utilizador com uma possível resposta, de modo a que quando uma correspondência é feita, a resposta seguinte mais próxima possa estar relacionada ao assunto dado e talvez seja uma correspondência possível.

O quinto e último critério consiste no contexto da interação atual, ou seja, compara os pares perguntas e respostas anteriores no diálogo entre o utilizador e o SSS.

Os cinco critérios são importantes para o funcionamento do Filipe, porém o PALbot apenas necessitará dos três primeiros critérios.

3.4.2. Configurações do *Say Something Smart*

O Filipe, como referido anteriormente no capítulo 2, é um *chatbot* que utiliza o algoritmo SSS e que trabalha com a língua inglesa. A alteração para a língua portuguesa é ilustrada na Figura 33, nas linhas 4 e 5 do ficheiro `config.txt`.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <config>
3
4 <!-- Language: english or portuguese-->
5 <language>portuguese</language>
6
7 <!-- Lucene options -->
8 <.lucene>
9 <dbPath>/resources/luceneDB/debug/english/db.db4o</dbPath>
10 <indexPath>./resources/luceneIndexes/</indexPath>
11 <usePreviouslyCreatedIndex>false</usePreviouslyCreatedIndex>
12 <hitsPerQuery>5</hitsPerQuery>
13 </.lucene>

```

Figura 33 – Campo de alteração de linguagem no ficheiro `config.txt` do SSS.

No ficheiro configuração foram ainda alteradas diversas propriedades, de modo a se adaptarem as mesmas às necessidades do *chatbot* a desenvolver. Entre elas, a lista de *stopwords*⁴⁰ e de *normalizers*⁴¹ (Figura 34) – tendo em conta a língua escolhida no passo anterior –, e as respostas por omissão – devolvidas pelo *chatbot*, de modo aleatório, quando não encontra nenhum candidato, como se pode observar nas linhas 57 e 58 da Figura 35.

```

26 <!-- Normalizers: RemoveDiacriticalMarks, EnglishLemmatizer, PortugueseStemmer -->
27 <!--<normalizers>PortugueseStemmer</normalizers>-->
28
29 <normalizers>PortugueseStemmer,RemoveDiacriticalMarks</normalizers>

```

Figura 34 – Configuração de *normalizers* e *stopwords* no SSS.

```

56 </noAnswerFound>
57 <msg>Peco desculpa, não percebi</msg>
58 <msg>Não sei responder a isso</msg>
59 </noAnswerFound>

```

Figura 35 – Respostas de fora de domínio.

Inicialmente o SSS está configurado para selecionar a resposta com maior pontuação entre 20 respostas candidatas. Para este projeto, o número de respostas candidatas foi adaptado de forma a encontrar valores mais plausíveis, uma vez que a maior parte das 20 candidatas iniciais tinham uma pontuação nula. Este valor foi inicialmente reduzido para metade e por último para cinco respostas candidatas, como se pode observar na linha 12 da Figura 33, onde a pontuação das respostas perto de zero já era aceitável.

3.5. Comparação entre as *frameworks* consideradas

A Tabela 2 apresenta as principais diferenças e semelhanças entre as duas abordagens, refletindo as fases de adaptação do código, base de conhecimento, manutenção e utilização.

De acordo com a Tabela 2, a ALICE está disponibilizada gratuitamente e é de *open-source*. A aquisição do programa teve como comparar o mesmo com o SSS que foi disponibilizado pelo L2F para perceber qual dos dois algoritmos é o melhor para desenvolver o PALbot.

⁴⁰ Palavras que são filtradas.

⁴¹ Algoritmos que são implementados para omitir pontuações nos *inputs* dos utilizadores.

Tabela 2 – Comparação entre a ALICE e o SSS de acordo com as fases e subfases

Categorias		Algoritmos	ALICE	SSS
Fase	Subfase			
Adaptação do código	Fonte	<i>Open-source</i>		
	Funcionamento	Funciona por marcação de AIML	Funciona por medidas de similaridade onde as respostas com a pontuação mais alta são escolhidas.	
	Linguagem Natural	É necessário alterar manualmente para a linguagem portuguesa, inserindo conjunto de palavras portuguesas.	Preparado para linguagem portuguesa.	
Base de conhecimento	Utilização pelo programador	Fácil de aprender		
	Definição de variáveis	É possível criar facilmente variáveis e usá-las (como te chamas)	Não é possível	
	Arquitetura do Corpus	Introdução de perguntas e respostas não é linear.	Não há limitações	
Manutenção	Atualização da base de conhecimento	Processo automatizado		
Utilização	Criatividade nas respostas	Pode reagir com as perguntas do utilizador	Não reage às perguntas do utilizador, limita-se a responder às perguntas.	
	Assertividade	Resposta não será a desejada se a marcação falhar.	Ao lidar com as medidas de similaridade garante que a pergunta pode ser diferente com a que está na base de conhecimento	
	OOD	Não ter OOD determinado, ou seja, não consegue pesquisar fora do âmbito		

Como referido no capítulo 2, a ALICE utiliza o AIML como interpretador, enquanto o SSS recorre a medidas de similaridade, que definem a resposta para o *input* do utilizador.

O SSS já vem preparado para lidar com palavras da Língua Portuguesa e tem preocupação com *stopwords*, enquanto que no AIML é necessário fazer toda a preparação manual no algoritmo.

Como referido acima, o *corpus* foi processado de forma automática para agilizar a atualização do ficheiro que contém a base de conhecimento dos *chatbots*. Assim, se for necessário retirar ou inserir perguntas e respostas, bastará eliminar ou adicionar linhas no Excel e voltar a gerar a base de conhecimento.

Uma grande vantagem do AIML é a criação fácil de variáveis, como é referido no capítulo anterior. Ao usar o AIML, é possível reagir com perguntas do utilizador, uma vez que o AIML tem a capacidade de personalizar perguntas e respostas. Por outro lado, o SSS, ao lidar com medidas de similaridade, garante que cada entrada do utilizador é correspondida na base de conhecimento, se esse assunto estiver presente na mesma.

Para a introdução de perguntas e respostas na ALICE é necessária uma adaptação destas para frases curtas, uma vez que o conteúdo do *pattern* (conteúdo da pergunta) tem de ter uma curta frase para que haja uma maior probabilidade de correspondência com o *input* do utilizador. No SSS, é somente necessário corresponder perguntas e respostas com o I e o R respetivamente.

A principal desvantagem do AIML reside na necessidade de o utilizador escrever sempre a marcação que está na base de dados para ter a resposta desejada.

No AIML existe a preocupação na forma em que o utilizador pode fazer a pergunta, e utiliza-se o *wildcard* “*” para colmatar situações de palavras que possam estar antes e/ou depois do *pattern* definido. Sendo assim, é necessário definir com certo cuidado o que vai estar no *pattern* para que haja correspondência com os *inputs* do utilizador. No SSS, por sua vez basta que exista no *input* do utilizador uma parte em comum com alguma pergunta da base de conhecimento, embora esse *input* possa levar a respostas erradas.

Ambos os algoritmos têm a mesma estratégia para lidar com interações OOD, ou seja, lidam da mesma forma com um conjunto de mensagens pré-definidas caso não haja correspondência de perguntas e respostas.

Posto isto, o melhor será utilizar o SSS, uma vez que esta lida bem com a pluralidade das palavras e assim não existe a preocupação de saber como o utilizador irá fazer a sua pergunta, embora o AIML seja mais interativo. Tratando-se também de uma ferramenta para esclarecimento de dúvidas da parte do utilizador seria mais viável usar o SSS de modo a garantir que se o *input* do utilizador estiver na base de conhecimento, o *chatbot* devolverá a resposta correta.

3.6. Controlo da base de conhecimento

Após concluir que o PALbot será desenvolvido com base no SSS, será necessário elaborar um método de controlo, sobre a viabilidade das respostas, ou seja, controlar as respostas devolvidas pelo PALbot, face às perguntas dos utilizadores.

De forma a haver um maior controlo sobre as respostas, como se pode observar na Figura 36, é necessário desenvolver processos automatizados.

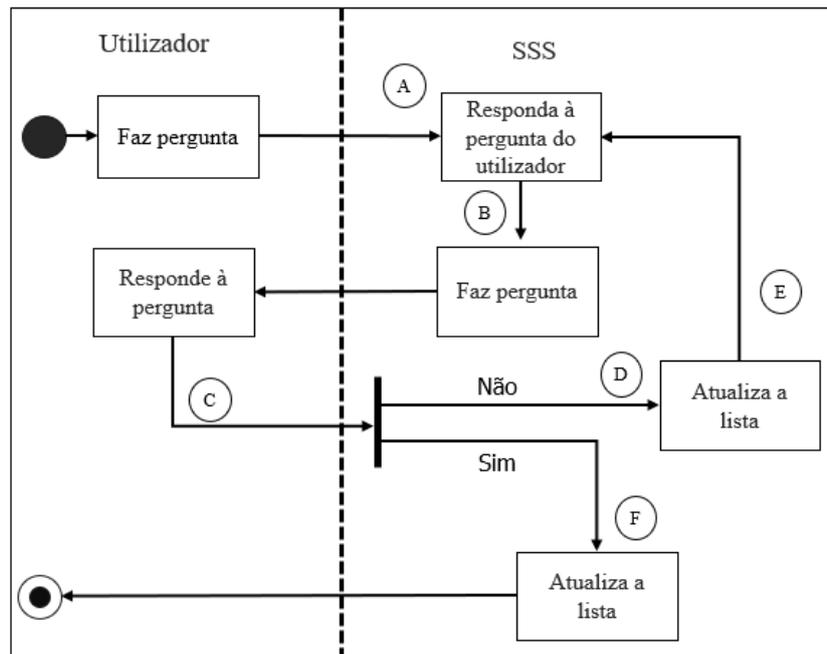


Figura 36 – Diagrama de atividades sobre a apreciação das respostas pelo utilizador.

De acordo com a Figura 36, o diagrama começa com o processo A onde o SSS responde à pergunta do utilizador (linhas 1 e 2 da Figura 37). Em seguida, no processo B, o SSS pergunta ao utilizador se a informação é útil, sendo que, no processo C, o mesmo devolve a resposta de sim ou não (linhas 3 e 4 da Figura 37).

```

1 Utilizador: quando se escreve à ou há?
2 PALbot: À é a contracção da preposição a com o artigo definido no feminino singular a. Ex.
3 PALbot: Esta informação foi útil : (Sim/Nao)
4 Utilizador:

```

Figura 37 – Diálogo entre o utilizador e o PALbot para a confirmação da viabilidade da informação.

Os processos D e E da Figura 36 refletem a resposta do utilizador como “não” (linha 4 da Figura 38). Uma vez que o SSS tem cinco possíveis candidatos, quando o utilizador responde que a informação não é útil, o SSS devolve o segundo candidato com maior pontuação e assim sucessivamente. Caso o último candidato não tenha sido uma

informação útil, o SSS responde com uma resposta pré-definida, como se pode observar na linha 17.

```

1 Utilizador: quando se escreve à ou há?
2 PALbot: À é a contracção da preposição a com o artigo definido no feminino singular a. Ex.
3 PALbot: Esta informação foi útil : (Sim/Não)
4 Utilizador: Não
5 PALbot: l - Escreve-se porquea Quando é conjunção causal Não saio, porque está a chover.b Qu
6 PALbot: Esta informação foi útil : (Sim/Não)
7 Utilizador: Não
8 PALbot: Por extenso, a expressão 25% escreve-se da seguinte forma vinte e cinco por cento, e
9 PALbot: Esta informação foi útil : (Sim/Não)
10 Utilizador: Não
11 PALbot: No presente do indicativo, os verbos de tema em a ou seja, os acabados em -ar, no in
12 PALbot: Esta informação foi útil : (Sim/Não)
13 Utilizador: Não
14 PALbot: Embora sejamos iguais perante a natureza, não o somos perante a sociedade, porque ne
15 PALbot: Esta informação foi útil : (Sim/Não)
16 Utilizador: Não
17 PALbot: A sua resposta não foi encontrada.
18 Utilizador:

```

Figura 38 – Diálogo entre o utilizador e o PALbot caso a resposta seja incorreta.

O processo F da Figura 36, é iniciado quando o utilizador confirma a viabilidade da resposta, como se pode observar na Figura 39.

```

1 Utilizador: quando se escreve à ou há?
2 PALbot: À é a contracção da preposição a com o artigo definido no feminino singular a. Ex.
3 PALbot: Esta informação foi útil : (Sim/Não)
4 Utilizador: Sim
5 PALbot: Em que mais posso ajudar?
6 Utilizador:

```

Figura 39 – Diálogo entre o utilizador e o PALbot caso a resposta seja correta.

Em ambos os casos, à medida que o utilizador vai respondendo à pergunta do SSS, a lista do controlo de informação vai sendo atualizada nos processos D e F da Figura 36.

O controlo de informação tem o objetivo entender se as respostas são corretas para as perguntas solicitadas, e consiste na atualização de uma tabela no ficheiro InformationControl, onde estão contidas informações de respostas Sim/Não relativamente aos respetivos IDCiberduvidas e ID como se pode observar na Tabela 3.

A Tabela 3 tem cinco colunas em que o *input* do utilizador gera cinco candidatos e o mesmo considera que nenhuma das respostas está certa.

Tabela 3 – Tabela com informações caso as respostas sejam incorretas.

IDCiberduvidas	ID	Sim	Não	Total
7411	0	0	1	1
7411	1	0	1	1
30195	0	0	1	1
4345	0	0	1	1
32589	1	0	1	1

Na Tabela 4, o *input* do utilizador gera uma resposta e o mesmo valida a resposta.

Tabela 4 – Tabela com informações caso o primeiro candidato seja resposta correta.

IDCiberduvidas	ID	Sim	Não	Total
11238	0	1	0	1

Na Tabela 5, o *input* do utilizador gera uma resposta, mas o utilizador só valida a resposta como “Sim” à terceira vez.

Tabela 5 – Tabela com informações caso o terceiro candidato seja correto.

IDCiberduvidas	ID	Sim	Não	Total
10638	0	0	1	1
10638	1	0	1	1
24498	0	1	0	1

Como referido, o InformationControl é elaborado de forma a controlar o fluxo da informação. Assim, o parâmetro “Sim” é formatado de forma a ficar de cor verde quando a contagem atinge valores acima de dez, e o parâmetro “Não” é configurado para ficar vermelho quando a contagem ultrapassa do valor cinco, como se pode observar na Tabela 6.

Tabela 6 – Tabela com informações mais complementares.

IDCiberduvidas	ID	Sim	Não	Total
10638	0	12	1	13
10638	1	2	10	12
24498	0	17	6	23
30126	0	0	1	1
30126	1	1	0	1
23680	0	15	4	19
23680	1	0	1	1
25133	1	2	3	5
25133	0	1	0	1

Assim, o InformationControl poderá melhorar a base de conhecimento do PALbot com as opções que o utilizador poderá responder como correta ou incorreta, ajudando o gestor de informação a atualizar o conhecimento do agente de conversação.

3.7. Aprendizagem

Uma outra funcionalidade implementada no PALbot é a aprendizagem. Esta funcionalidade consiste em aumentar automaticamente a capacidade do *chatbot* de responder a novas formulações de perguntas sobre a mesma base de conhecimento de novas sugestões.

Primeiramente, o PALbot dá a resposta mediante a pergunta do utilizador. De seguida, o mesmo pergunta ao utilizador sobre a correção da resposta fornecida. Se o utilizador validar a resposta, então o PALbot irá medir a distância da pergunta do utilizador com

todas as perguntas que estão na base de conhecimento. Quando essa distância for superior a 10% e inferior a 80%, então o PALbot guardará a pergunta e a resposta na sua base de conhecimento. O intervalo é determinado de forma que as perguntas tenham proximidade entre elas para garantir a validade das respostas. Caso o utilizador conteste a validade da resposta fornecida, o PALbot fornecerá uma nova resposta e este ciclo irá repetir até esgotar os cinco candidatos.

A Figura 40 apresenta o par pergunta e resposta na base de conhecimento e a Figura 41 demonstra uma das maneiras em que a pergunta da Figura 40 pode ser feita.

```
SubID - 0
DialogId - 0
Diff - 0
IDCiberduvidas - 10346
ID - 0
I - História e estória
R - O Aurélio - Novo Dicionário da Língua Portuguesa, quando procuramos a palavra estória,
```

Figura 40 – Conjunto de pergunta e resposta da base de conhecimento.

```
1 Utilizador: Qual é a diferença entre estória e história?
2 PALbot: O Aurélio - Novo Dicionário da Língua Portuguesa, quando procuramos a palavra estória,
3 PALbot: Esta informação foi útil : (Sim/Não)
4 Utilizador: sim
5 PALbot: Em que mais posso ajudar?
6 Utilizador:
```

Figura 41 – Diálogo entre o PALbot e o utilizador.

O controlo de aprendizagem que o PALbot possuirá encontra-se exemplificado na Tabela 7. Na primeira coluna é apresentado o *input* do utilizador e a segunda coluna é caracterizada pelo IDCiberdúvidas da resposta que é fornecida. Posto isto, a Tabela 7 controla a viabilidade das respostas com as novas perguntas introduzidas na base de conhecimento onde a coluna “Sim” é marcada com cor verde acima de duas contagens e a coluna “Não” assinala a vermelho acima de uma contagem. Este controlo de aprendizagem está situado no InformationControl que é um módulo responsável para guardar informação sobre correções das respostas às perguntas. Uma vez que a linha 1 da Figura 41 cumpre os requisitos de distância com a pergunta que está na base de conhecimento, a mesma será inserida no PALbot resultando a inserção dos dados que estão exemplificados na Tabela 7.

Tabela 7 – Controlo da aprendizagem do PALbot.

Pergunta	IDCiberdúvidas da resposta	Sim	Não	Total
Qual é a diferença entre estória e história?	10346	5	3	8
Como se pronuncia o plural de ufir?	14784	1	1	2
Qual é o significado de XPTO?	14774	6	1	7
Como se chama o conjunto de soldados?	32721	0	1	1
A palavra bidiário existe?	16145	0	7	7
O verbo postar de post existe?	16139	1	1	2

Capítulo 4 – Avaliação e testes

Neste capítulo, é avaliado o sistema desenvolvido a partir de duas estratégias, a avaliação da correção das respostas e a avaliação do desempenho na perspetiva do utilizador.

A avaliação da correção das respostas consiste em perceber quantas respostas corretas e incorretas são fornecidas pelo PALbot e pelo motor de pesquisa do Ciberdúvidas tendo sido usada para avaliar a métrica *Mean Reciprocal Rank*⁴² (MRR).

Para avaliar o desempenho na perspetiva do utilizador elaborou-se um inquérito a partir do Google *Forms*⁴³, para uma amostra aleatória, com o objetivo de inferir o nível de acessibilidade que o PALbot e o motor de pesquisa do Ciberdúvidas poderão ter. Adicionalmente, foram executados testes de forma a avaliar o desempenho da ferramenta.

4.1. Apresentação do estudo e resultados obtidos

Os dados do inquérito foram recolhidos presencialmente, uma vez que o PALbot apresenta um ficheiro demasiado pesado para a disponibilização do mesmo em formato *online*. O período de recolha dos dados ocorreu entre os dias 7 e 21 de maio de 2019. Para o tratamento estatístico dos dados recolhidos utilizou-se o suporte informático SPSS e Excel, que permitiram a elaboração de gráficos e tabelas apresentados.

O inquérito realizado encontra-se em anexo (Anexo 1) e as 13 perguntas elaboradas juntamente com o seu respetivo objetivo apresentam-se especificadas na Tabela 8.

Denota-se que nas perguntas de resposta aberta, nomeadamente “Dê a sua opinião”, é difícil quantificar os resultados, uma vez que não era de carácter obrigatório para os inquiridos responderem e a amostra não foi significativamente o suficiente, tendo sido descartadas da avaliação.

⁴² Avaliação baseada em recuperação do diálogo, na qual o primeiro candidato relevante é recuperado.

⁴³ Plataforma gratuita para elaboração, disponibilização e preenchimento de inquéritos.

Tabela 8 – Caracterização dos objetivos das perguntas elaboradas no inquérito.

	Perguntas do questionário	Objetivos das perguntas
Características sociodemográficas e indicador de conhecimento do <i>chatbot</i>	Sexo	Identificação das características sociodemográficas.
	Idade	
	Habilitações Literárias	
	Indique a sua área de formação	
	Já conversou alguma vez com um <i>chatbot</i> ?	Análise da amostra que já tiveram contacto com um <i>chatbot</i> .
Interação com o <i>chatbot</i> - avaliação controlada	Obteve a informação que procurava por parte do <i>chatbot</i> ?	Avaliação do desempenho, eficácia do <i>chatbot</i> com o motor de pesquisa do Ciberdúvidas. A avaliação controlada baseia-se em perguntas que o utilizador coloca em função de um conjunto de tópicos que é fornecido retirados dos dados fornecidos pela plataforma Ciberdúvidas
	Obteve a informação que procurava por parte do Ciberdúvidas?	
	Dê a sua opinião	Entender qual é opinião dos utilizadores no âmbito das respostas que são devolvidas na avaliação controlada
Interação com o <i>chatbot</i> - avaliação livre	Obteve a informação que procurava por parte do <i>chatbot</i> ?	Avaliação do desempenho, eficácia e eficiência percebida do <i>chatbot</i> com o motor de pesquisa do Ciberdúvidas. A avaliação não controlada baseia-se em qualquer pergunta que o utilizador coloca na matéria da Língua Portuguesa.
	Obteve a informação que procurava por parte do Ciberdúvidas?	
	Dê a sua opinião	Entender qual é opinião dos utilizadores no âmbito das respostas que são devolvidas na avaliação controlada
Avaliação da eficiência percebida	O <i>chatbot</i> facilita o acesso à informação.	Analisar se os utilizadores reconhecem o <i>chatbot</i> como uma boa ferramenta para o acesso à informação.
	Dê a sua opinião	Entender qual é o nível de conforto que as pessoas estão em relação à utilização de um <i>chatbot</i> .

Para avaliação da interação com o PALbot num ambiente controlado, foi fornecida uma lista com tópicos perguntas e respostas para que os utilizadores elaborassem perguntas para introduzir no *chatbot* a partir dos tópicos escolhidos. Por sua vez, para avaliação da interação em ambiente livre, as perguntas foram colocadas livremente – sem

qualquer tópico de apoio – desde que as mesmas estivessem relacionadas com dúvidas sobre a Língua Portuguesa.

4.1.1. Caracterização da amostra

Uma vez terminado o processo de recolha de dados, através da realização do inquérito a uma amostra total composta por 73 indivíduos, inseriram-se os dados obtidos na ferramenta SPSS, na qual se definiu a variável idade segundo quatro gerações (Berkup, 2014)⁴⁴ apresentadas na Tabela 9.

Tabela 9 – Caracterização do intervalo de idades com as gerações.

	Boomer	Geração X	Geração Y	Geração Z
Intervalo de ano de nascimento	≤ 1664	1965-1979	1980-1994	≥ 1995
Intervalo de idades	≥ 55 anos	40 - 54 anos	25 - 39 anos	≤ 24 anos

As variáveis género, idade, habilitações literárias e área de formação foram recolhidas aquando da realização do inquérito, permitindo a caracterização sociodemográfica dos inquiridos, apresentada na Tabela 10.

Tabela 10 – Caracterização sociodemográfica dos inquiridos.

Características		Inquiridos	Percentagem (%)
Idade	Geração X	7	9,59
	Geração Y	49	67,12
	Geração Z	16	21,92
	Boomer	1	1,37
	Total	73	100,00
Sexo	Feminino	29	39,73
	Masculino	44	60,27
	Total	73	100,00
Habilitações Literárias	12.º ano	22	30,14
	Curso Tecnológico/Profissional/Outros (Nível III)	1	1,37
	Licenciatura	20	27,40
	Mestrado	25	34,25
	Doutoramento	5	6,85
	Total	73	100,00

⁴⁴ Working With Generations X And Y In Generation Z Period: Management Of Different Generations In Business Life.

Área de formação	Artes	5	6,85
	Ciências	26	35,62
	Economia	8	10,96
	Engenharias	18	24,66
	Humanidades	16	21,92
	Total	73	100,00

Relativamente ao género, pode observar-se que a maioria dos inquiridos são do sexo masculino (60,27%). Quanto à idade, a maior parte dos inquiridos integram a geração Y, com idades compreendidas entre 25 e 39 anos. No que diz respeito a habilitações literárias e área de formação, verifica-se que a maioria dos inquiridos possui grau de mestre (34,25%) e uma formação na área de ciências (31,51%). Tendo em conta que na geração Boomer foi apenas registado um inquirido, considerou-se que amostra não seria representativa desta geração, excluindo-a de análise.

4.1.2. Avaliação em ambiente controlado

A avaliação em ambiente controlado indica que o utilizador testará o PALbot com perguntas relacionadas com os tópicos fornecidos, de forma a entender se o *chatbot* fornece respostas com perguntas diferentes para o mesmo tópico.

A avaliação genérica em ambiente controlado do PALbot e do Ciberdúvidas, apresentada na Figura 42, demonstra o parecer dos utilizadores mediante o desempenho das duas ferramentas. No PALbot, observam-se maiores percentagens nas opções de resposta “Geralmente”, “Quase sempre” e “Sempre”, com percentagens de 34,20%, 37,00% e 17,80%, respetivamente.

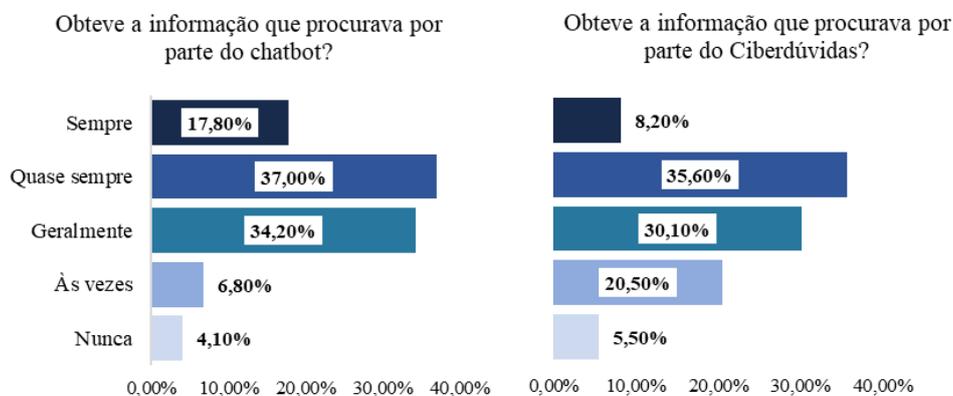


Figura 42 – Indicadores estatísticos da avaliação geral do PALbot e do Ciberdúvidas em ambiente controlado.

No Ciberdúvidas, as mesmas opções de resposta são também elevadas, mas inferiores às registadas para o PALbot, nomeadamente no caso da opção “Sempre” com aproximadamente metade das respostas, quando comparado com o agente de conversação. Consequentemente, na plataforma, os parâmetros “Às vezes” e “Nunca” apresentam valores superiores em relação à outra ferramenta.

A avaliação em ambiente controlado relativos às três gerações do PALbot e do Ciberdúvidas é apresentada na Figura 43.

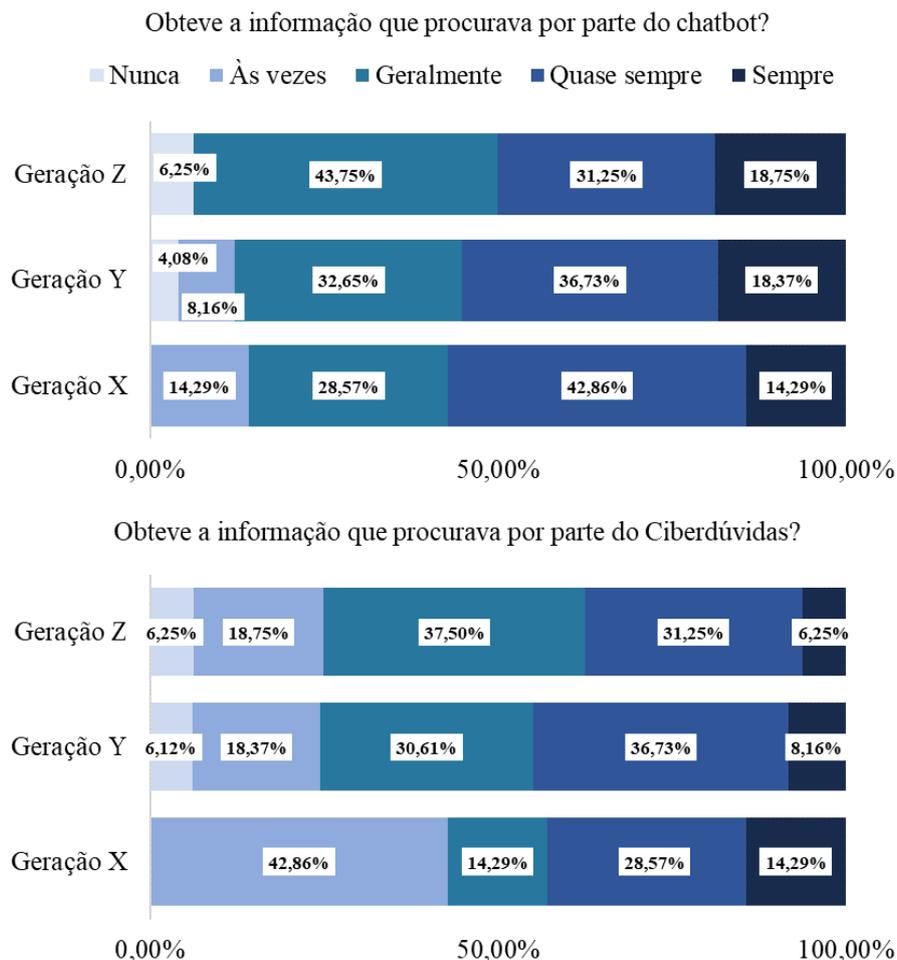


Figura 43 – Indicadores estatísticos das gerações em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente controlado.

Observa-se que, para o PALbot, as opções de resposta consideradas positivas - “Quase sempre” e “Sempre” – juntamente com a opção “Geralmente” são as mais expressivas transversalmente às três gerações, com destaque para as opções “Quase sempre” e “Geralmente”. Por sua vez, no caso do motor de pesquisa do Ciberdúvidas, as mesmas opções positivas, bem como “Geralmente” perdem expressão para as opções consideradas negativas, “Às vezes” e “Nunca”, com destaque para a primeira. Deve ainda salientar-se

que, para a geração X, em particular, a opção negativa “Às vezes” regista a percentagem mais elevada, com mais de 40% das respostas.

Os gráficos de avaliação em ambiente controlado relativos a habilitações literárias, são ilustrados na Figura 44.

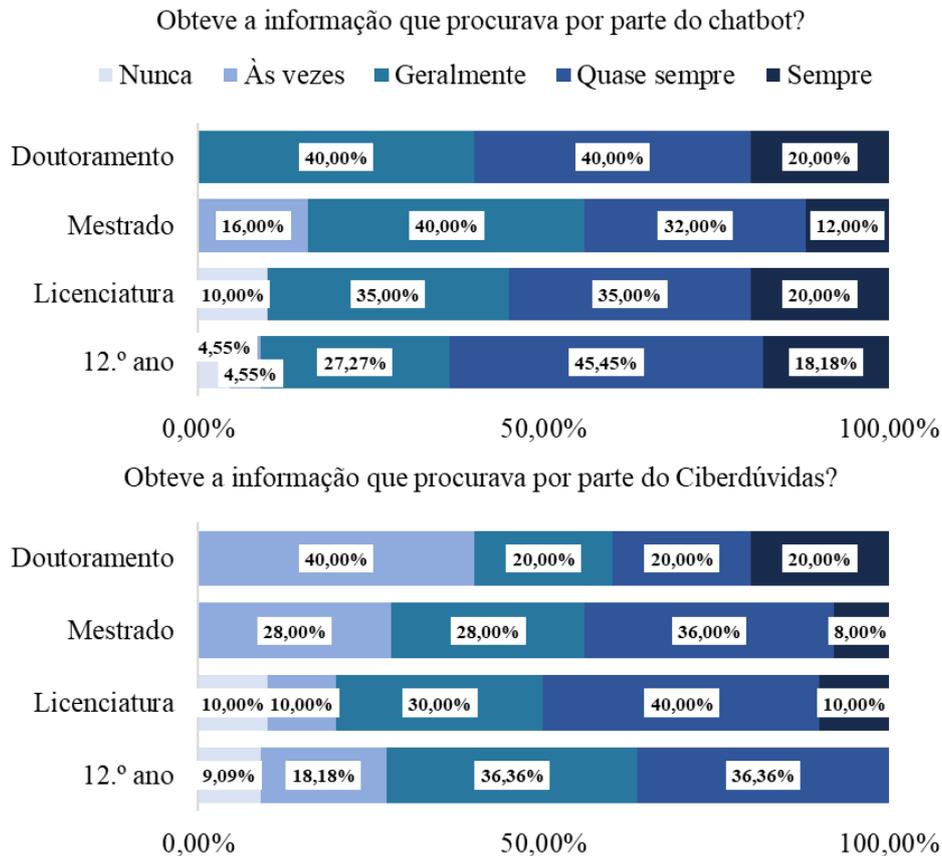


Figura 44 – Indicadores estatísticos das habilitações literárias em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente controlado.

É possível constatar que transversalmente a todas as diferentes habilitações literárias, as percentagens das opções consideradas como positivas são mais elevadas no PALbot comparativamente ao motor de pesquisa do Ciberdúvidas. No PALbot, a opção “Geralmente” é mais expressiva em todas as habilitações, exceto para inquiridos com o 12.º ano. Relativamente a opções de resposta negativas, as mesmas encontram-se ausentes na avaliação do *chatbot* por inquiridos com o grau de doutoramento, e têm valores poucos expressivos tanto em inquiridos com mestrado (16,00%), licenciatura (10,00%) ou 12.º ano (5,10%). Porém, no Ciberdúvidas, estas mesmas opções registaram resultados significativos no grau de doutoramento com 40,00%, no de mestrado com 28,00%, no de licenciatura com 20,00% e no de 12.º ano com 27,27%.

Quanto aos gráficos relativos às áreas de formação em ambiente controlado encontram-se apresentados na Figura 45.

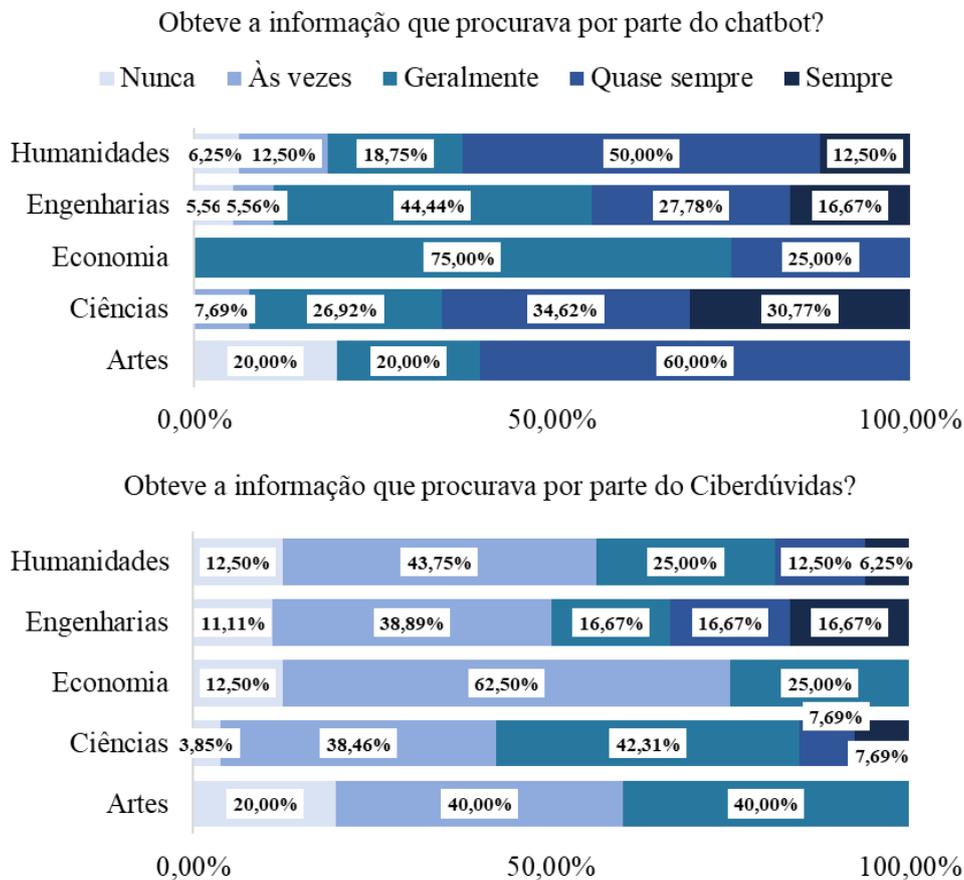


Figura 45 – Indicadores estatísticos das áreas de formação em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente controlado.

Observa-se que, na avaliação da qualidade da informação dada pelo *chatbot*, as opções de resposta positivas, consideradas anteriormente, surgem nas áreas de humanidades, engenharias e ciências, sendo que nas áreas de economia e artes, apenas se encontra presente a opção de resposta “Quase sempre”. Em conjunto, as duas opções positivas têm o valor mais alto na área de ciências com 65,39% e humanidades com 62,50%. Relativamente às opções de resposta negativa, estas são mais expressivas na área de artes, onde a opção “Nunca” prevalece com uma percentagem superior em comparação com as restantes áreas, contabilizando 20,00% das respostas. A opção “Geralmente” apresenta uma maior percentagem na área de economia com 75,00%.

No caso do motor de pesquisa do Ciberdúvidas, as opções “Quase sempre” e “Sempre” têm, igualmente, em conjunto, o valor mais alto nas áreas de engenharias com 33,34% e humanidades com 18,75%. No entanto, estes valores são muito mais baixos

comparativamente aos mesmos obtidos no PALbot. As opções “Nunca” e “Às vezes” têm o valor mais alto na área de economia com 75% e a opção “Geralmente” é predominante na área de ciências com 42,31%.

4.1.3. Avaliação em ambiente livre

A experimentação em ambiente livre indica que o utilizador poderá fazer qualquer tipo de pergunta relativamente a dúvidas sobre a Língua Portuguesa, avaliando se o PALbot consegue responder corretamente a este tipo de perguntas.

A avaliação genérica em ambiente livre do PALbot e do Ciberdúvidas, representada na Figura 46, apresenta o parecer dos utilizadores ao testarem estas duas ferramentas sem qualquer apoio de tópicos. Constata-se que, no caso do PALbot, as opções “Às vezes” e “Geralmente” são as mais elevadas com 42,50%, 30,10%, respetivamente. No caso do motor de pesquisa do Ciberdúvidas, as opções referidas anteriormente são, também, as mais expressivas. No entanto, a opção “Nunca” regista a terceira maior percentagem (23,30%), mais de metade da percentagem correspondente obtida para o *chatbot*.

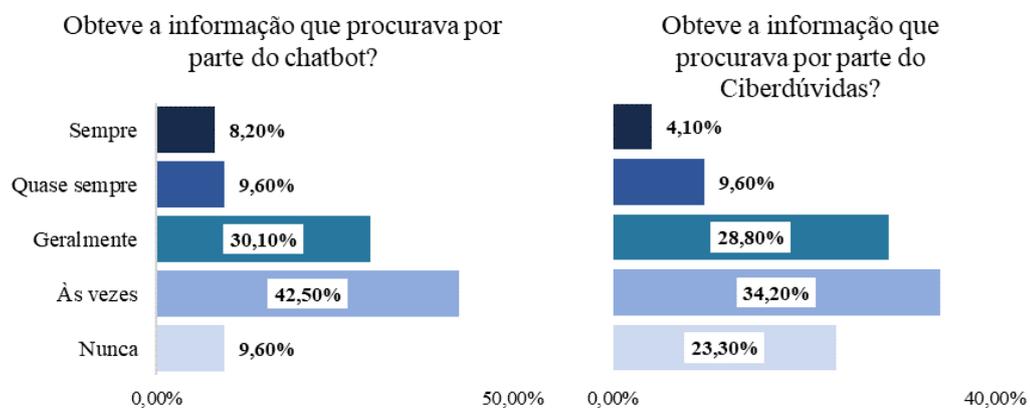


Figura 46 – Indicadores estatísticos da avaliação geral do PALbot e do Ciberdúvidas em ambiente livre.

Pelos gráficos da avaliação em ambiente livre relativos às três gerações, ilustrados na Figura 47, constata-se que para o PALbot as opções “Geralmente” e “Às vezes” são, transversalmente, as mais expressivas, com destaque para a opção “Às vezes” na geração X que contabiliza uma percentagem superior a 50,00%. Relativamente ao motor de pesquisa do Ciberdúvidas, às opções suprarreferidas junta-se a opção negativa “Nunca” com percentagens bastante mais elevadas, comparativamente ao PALbot, sobretudo no caso das gerações Z e Y, com percentagens de, aproximadamente, 25,00%.

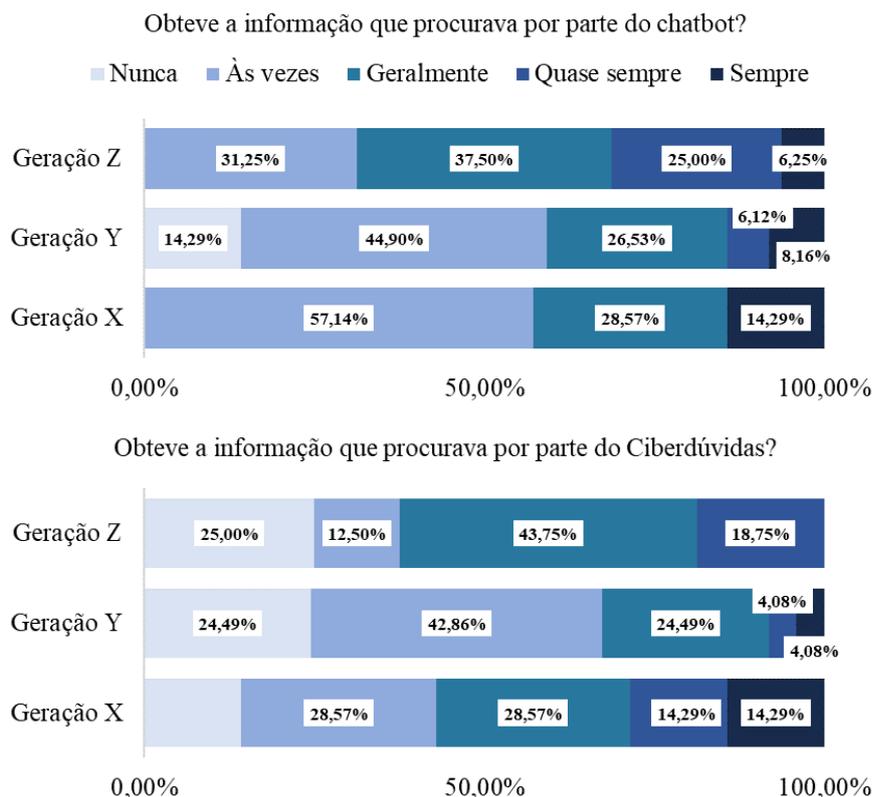


Figura 47 – Indicadores estatísticos das gerações em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente livre.

Os gráficos de avaliação em ambiente livre relativos a habilitações literárias, apresentados na Figura 48, permitem constatar que, no caso do PALbot, as opções “Às vezes” e “Geralmente” são as que apresentam, transversalmente, percentagens mais elevadas. A opção “Nunca” surge, apenas, no caso de inquiridos com Licenciatura (20,00%) ou 12.º ano (13,64%). As opções de resposta positivas são mais expressivas para o grau de Doutoramento, o que também se verifica para o motor de pesquisa do Ciberdúvidas. Contudo, neste caso, a opção “Às vezes” perde expressão relativamente às percentagens registadas com o PALbot e ganham expressão as percentagens relativas à opção negativa “Nunca”. Este último regista a percentagem mais elevada para inquiridos com o 12.º ano (36,36%).

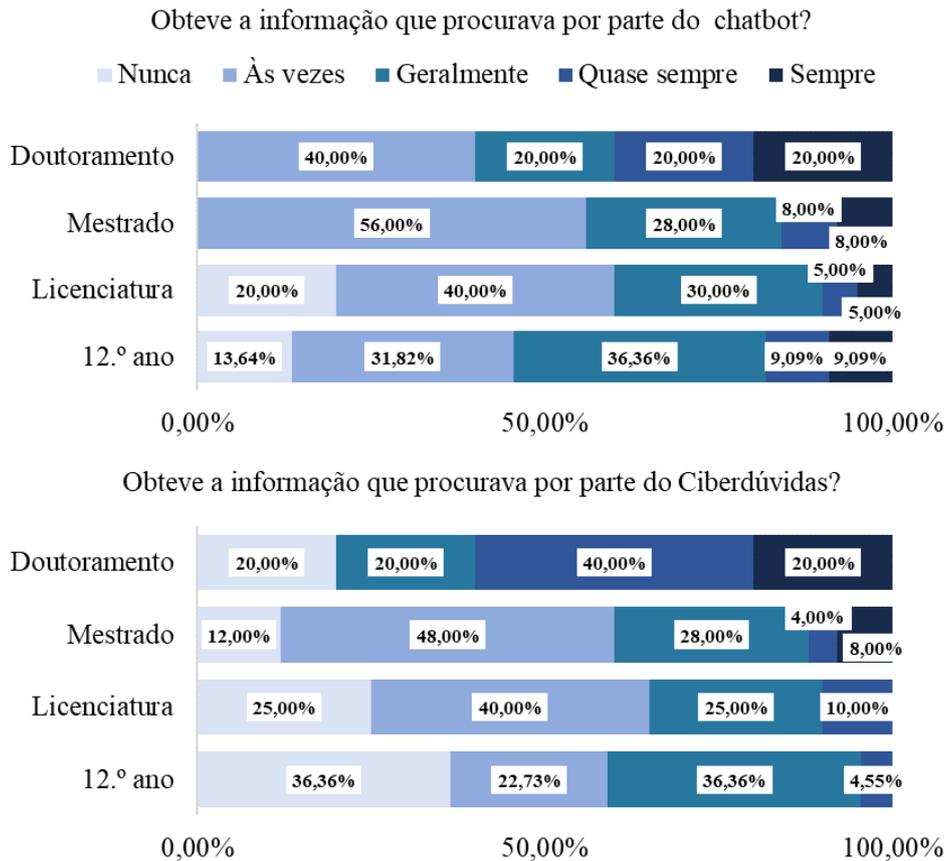


Figura 48 – Indicadores estatísticos das habilitações literárias em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente livre.

Pelos gráficos de avaliação em ambiente livre relativos a áreas de formação, ilustrados na Figura 49, verifica-se que, na avaliação da qualidade da informação dada pelo PALbot, as opções “Geralmente” e “Às vezes” são, transversalmente, as mais expressivas. As opções de resposta positivas consideradas, apenas estão presentes nas áreas de engenharias, humanidades e ciências, com destaque para a primeira onde estas opções perfazem um total de 33,34%. A opção “Nunca” assume a percentagem mais elevada para as artes (20,00%). Quanto ao Ciberdúvidas, constata-se que as opções positivas, em comparação com o caso anterior, assumem menor expressão para as áreas de engenharias, humanidades e ciências. Contudo, a opção “Quase sempre” regista respostas nas áreas de economia e das artes. Também, em todas as áreas, a opção negativa “Nunca” assume percentagens mais elevadas do que no PALbot, com destaque para as áreas de humanidades (50,00%) e economia (37,50%).

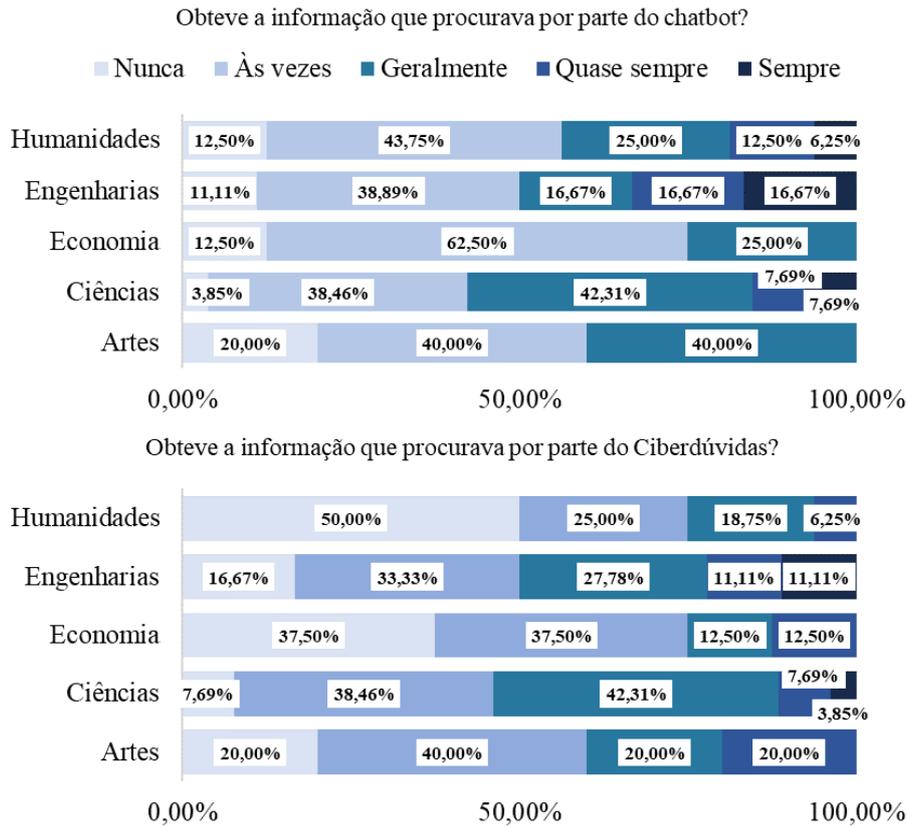


Figura 49 – Indicadores estatísticos das áreas de formação em relação à informação obtida por parte do PALbot e do Ciberdúvidas em ambiente livre.

4.1.4. Análise de relevância do PALbot

O histórico de interações com agentes de conversação encontra-se apresentado na Figura 50. Observa-se que a percentagem de inquiridos que afirma já ter experimentado uma conversação com um *chatbot* regista uma percentagem de 54,79%. A percentagem de inquiridos que nunca entraram em contacto com um agente de conversação é de 45,21%.

Já conversou alguma vez com um chatbot?

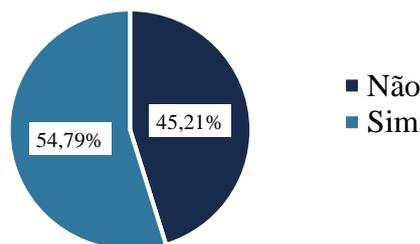


Figura 50 – Indicadores estatísticos do histórico de iterações com chatbots.

Relativamente ao grau de concordância com a afirmação “O *chatbot* facilita o acesso à informação.”, ilustrado no gráfico da Figura 51, observa-se que a maior expressão se centra nas opções de respostas positivas, respetivamente com 65,75%. A opção “Geralmente” e a opção de resposta negativa “Às vezes” assumem respetivamente com 21,92 % e 12,33%. Torna-se interessante salientar que não se registou nenhuma resposta da opção “Nunca”.

O chatbot facilita o acesso à informação.

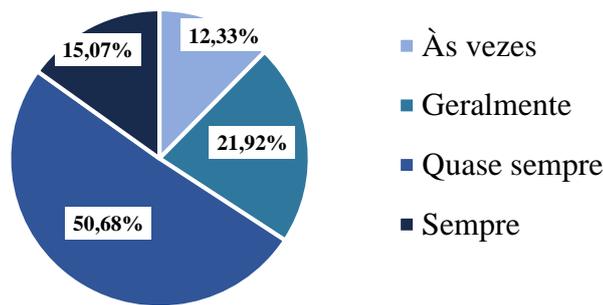


Figura 51 – Indicadores estatísticos da facilidade de acesso à informação fornecida pelo chatbot.

Na Figura 52 encontra-se representado um gráfico resultante do cruzamento de informação entre o histórico de conversação com um agente de conversação e o grau de concordância com a afirmação “O *chatbot* facilita o acesso à informação”.

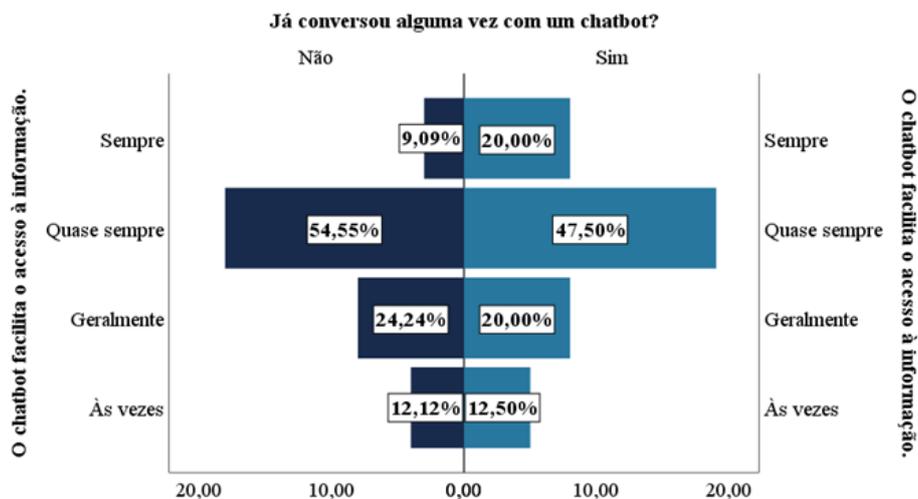


Figura 52 – Indicadores estatísticos do histórico de iterações com a facilidade do acesso à informação fornecida pelo chatbot.

Observa-se que no caso dos inquiridos que já tiveram contacto, no passado, com um agente de conversação, existe uma maior tendência para concordarem totalmente ou quase totalmente com a afirmação apresentada, comparativamente aos inquiridos que nunca tinham utilizado uma ferramenta deste género. No primeiro caso – inquiridos que responderam “Sim” regista-se uma concentração maioritária de respostas nas opções “Sempre” e “Quase sempre”. Por sua vez, no segundo caso, – inquiridos que responderam “Não” – observa-se que as percentagens mais expressivas correspondem às opções “Quase sempre” e “Geralmente”.

4.2. Avaliação do PALbot e do Ciberdúvidas

De forma a entender qual a eficácia do PALbot e do motor de pesquisa do Ciberdúvidas mediante solicitação de respostas, os mesmos foram avaliados utilizando, para tal, as perguntas realizadas pelos inquiridos.

Tabela 11 – Resultados das perguntas solicitadas mediante as respostas corretas e incorretas.

	Corretas	Incorretas	Total
Respostas	251	132	383
Total	65,54%	34,46%	100,00%

Com base na Tabela 11, o PALbot tem 383 iterações (consultar Anexo 2), conseguindo responder corretamente em 65,54% dos casos e incorretamente em 34,46% dos casos. A Tabela 12 é caracterizada com 65,54% dos casos em que as perguntas são respondidas corretamente. Esta tabela contém os indicadores dos cinco candidatos de resposta quando uma pergunta é elaborada. Com as 251 iterações obtém-se a resposta correta com o primeiro candidato em 77,69% dos casos, e os restantes em 22,31%. Ou seja, em grande parte desta avaliação a resposta correta centra-se no primeiro candidato.

Tabela 12 – Resultados da localização das respostas corretas nos candidatos no PALbot.

	1º candidato	2º candidato	3º candidato	4º candidato	5º candidato	Total
Respostas Corretas	195	19	11	8	18	251
Total	77,69%	7,57%	4,38%	3,19%	7,17%	100,00%

Para avaliar o desempenho do PALbot, com base nos resultados apresentados na Tabela 12, recorre-se ao método MRR. O MRR consiste numa avaliação baseada em recuperação do diálogo com o PALbot, para calcular a média do recíproco da

classificação, na qual o primeiro candidato relevante é recuperado. O MRR é calculado a partir da equação (1), onde n é o número de questões, i é o número individual do candidato e r_i é a posição recíproca da resposta correta (Abdul-Kader & Woods, 2017).

$$\text{MRR} = \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} \quad (1)$$

Com os dados da Tabela 12 e a equação (1), calculou-se o MRR do PALbot apresentando um valor de 0,8517, como se pode visualizar na Equação 2, indicando que o sistema de recuperação de informação é 85,17% relevante.

$$\text{MRR} = \frac{1}{251} \sum_{i=1}^{251} \frac{1}{\text{rank}_i} = \frac{(\frac{1}{1} \times 195) + (\frac{1}{2} \times 19) + (\frac{1}{3} \times 11) + (\frac{1}{4} \times 8) + (\frac{1}{5} \times 18)}{251} = 0,8517 \quad (2)$$

Para avaliar o desempenho do motor de pesquisa do Ciberdúvidas, recorreu-se às perguntas em que o PALbot forneceu respostas corretas (Tabela 11), em que os resultados estão apresentados na Tabela 13.

Tabela 13 – Resultados da localização das respostas corretas nos candidatos no motor de pesquisa do Ciberdúvidas

	1º candidato	2º candidato	3º candidato	4º candidato	5º candidato	6º candidato	8º candidato	13º candidato	Total
Respostas Corretas	85	47	51	20	23	17	6	2	251
Total	33,86%	18,73%	20,32%	7,97%	9,16%	6,77%	2,39%	0,80%	100,00%

Com base nos dados da Tabela 13 e a equação (1), calculou-se o MRR do motor de pesquisa do Ciberdúvidas apresentando um valor de 0,5531, como se pode observar na Equação 3, indicando que o sistema de recuperação de informação é 55,31% relevante.

$$\text{MRR} = \frac{1}{251} \sum_{i=1}^{251} \frac{1}{\text{rank}_i} = \frac{(\frac{1}{1} \times 85) + (\frac{1}{2} \times 47) + (\frac{1}{3} \times 51) + (\frac{1}{4} \times 20) + (\frac{1}{5} \times 23) + (\frac{1}{6} \times 17) + (\frac{1}{8} \times 6) + (\frac{1}{13} \times 2)}{251} = 0,5531 \quad (3)$$

Posto isto, o PALbot é um sistema de recuperação mais relevante em comparação com o motor de pesquisa do Ciberdúvidas.

Um outro método de avaliação tem como objetivo compreender a eficácia tanto do PALbot como o motor de pesquisa do Ciberdúvidas quando são apresentadas com as mesmas questões. Estas questões são utilizadas para testar no PALbot enquanto que no Ciberdúvidas utiliza-se apenas as palavras-chaves das mesmas.

Para a pesquisa de perguntas no Ciberdúvidas, utilizam-se pares de perguntas e respostas em que o PALbot tem a resposta correta no quinto candidato. Utilizam-se as perguntas em que a resposta correta é o quinto candidato, de forma a entender se o

Ciberdúvidas será também capaz de fornecer as respostas corretas mediante a entrada da sentença no motor de busca da plataforma.

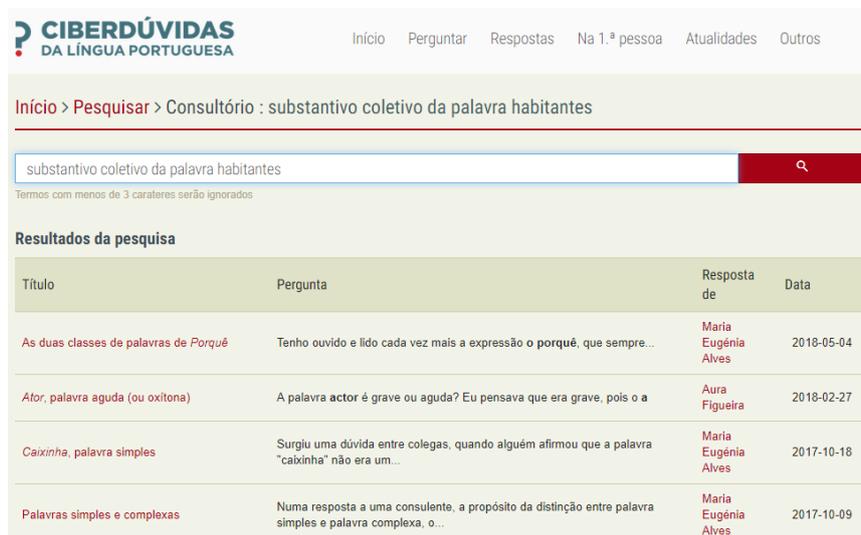
```

1 Utilizador: substantivo coletivo da palavra habitantes
2 PALbot: Se considerarmos que o substantivo natureza pode ser definido como conjunto de elementos do mu
3 PALbot: Esta informação foi útil : (Sim/Não)
4 Utilizador: nao
5 PALbot: Pode considerar-se - e muitas gramáticas assim o fazem - que trio é um numeral coletivo. O Dic:
6 PALbot: Esta informação foi útil : (Sim/Não)
7 Utilizador: nao
8 PALbot: Parece-me que o substantivo colectivo com o sentido de conjunto de habitantes será população. (
9 PALbot: Esta informação foi útil : (Sim/Não)
10 Utilizador: sim
11 PALbot: Em que mais posso ajudar?
12 Utilizador:

```

Figura 53 – Diálogo entre o utilizador e o PALbot.

A Figura 53 ilustra um exemplo de perguntas que o utilizador pode solicitar ao PALbot, em que o mesmo facilita respostas erradas. Um exemplo semelhante para a plataforma Ciberdúvidas, ilustrado na Figura 54, mostra que a mesma não reconhece as palavras chaves para indicar a resposta correta.



CIBERDÚVIDAS
DA LÍNGUA PORTUGUESA

Início Perguntar Respostas Na 1.ª pessoa Atualidades Outros

Início > Pesquisar > Consultório : substantivo coletivo da palavra habitantes

substantivo coletivo da palavra habitantes

Termos com menos de 3 caracteres serão ignorados

Resultados da pesquisa

Título	Pergunta	Resposta de	Data
As duas classes de palavras de <i>Porquê</i>	Tenho ouvido e lido cada vez mais a expressão o <i>porquê</i> , que sempre...	Maria Eugénia Alves	2018-05-04
<i>Ator</i> , palavra aguda (ou oxitona)	A palavra <i>actor</i> é grave ou aguda? Eu pensava que era grave, pois o a	Aura Figueira	2018-02-27
<i>Caixinha</i> , palavra simples	Surgiu uma dúvida entre colegas, quando alguém afirmou que a palavra "caixinha" não era um...	Maria Eugénia Alves	2017-10-18
Palavras simples e complexas	Numa resposta a uma consulente, a propósito da distinção entre palavra simples e palavra complexa, o...	Maria Eugénia Alves	2017-10-09

Figura 54 – Pesquisa na plataforma Ciberdúvidas (<https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=substantivo+coletivo+da+palavra+habitantes>).

A Tabela 14 (versão completa no Anexo 3) ilustra todas as perguntas (Anexo 2) em que as respostas corretas foram dadas no quinto candidato. Dessas perguntas foram retirados os tópicos para pesquisa no motor de pesquisa do Ciberdúvidas. Desta avaliação, baseada nos exemplos ilustrados nas Figuras 53 e 54, pode observar-se que o Ciberdúvidas não fornece todas as respostas, contrariamente ao que acontece com o PALbot.

Tabela 14 – Comparação dos resultados de pesquisa entre Ciberdúvidas e PALbot.

Pergunta	Ciberdúvidas
	Resposta correta
escreve-se em vez ou invés	Sim
escreve-se à partida ou apartida?	Sim
haja ou aja?	Sim
o que é um bisavô?	Sim
o que são homónimos?	Sim
palavra membra existe?	Não
precaver e prevenir é igual?	Não
qual a definição de amigo?	Sim
qual a forma correcta aonde ou onde?	Não
qual é o antónimo de bondade?	Não
qual é o sinónimo de envolvente?	Sim
Qual o mais correto preveram ou previram?	Não
quando se usa dar ou dá?	Não
quando uso hifens?	Não
viajem ou viagem?	Sim
qual deles o correto segmento ou seguimento?	Não
o que quer dizer Octeto	Sim
qual é o coletivo de avião?	Sim

Repetiu-se o mesmo método de avaliação na Tabela 15 (Anexo 4), porém todas as perguntas representadas (Anexo 2) foram selecionadas aleatoriamente nos casos em que o PALbot não encontrou respostas corretas. Observou-se que, no caso do Ciberdúvidas, apenas em 3 das 18 perguntas efetuadas foram dadas respostas corretas.

Tabela 15 – Comparação dos resultados de pesquisa entre Ciberdúvidas e PALbot.

Pergunta	Ciberdúvidas
	Resposta correta
diz se jurí ou júri?	Não
casa-se ou casasse	Não
haver e a ver?	Não
mais grande ou maior?	Sim
o que significa slb?	Não
à rua ou na rua?	Não
agradecer pela ou a	Não
Propício ou propenso?	Não
paralisação ou paralização?	Não
O que é um cão?	Não
a palavra quites existe?	Não
escreve-se ficção ou ficção?	Sim
significado de inócuo	Não
significado de desasnado	Não
Vírgula entre sujeito e predicado	Sim
significado facal	Não
Quando se deve usar um hiffen nas frases?	Não
ao meu ver ou a meu ver?	Não
supérgluo ou superfulo?	Não
significado de medeia?	Não

4.3. Discussão

O PALbot foi desenvolvido com o objetivo de ser utilizado por qualquer indivíduo da população que utiliza língua portuguesa. Posto isto, a representatividade da amostra em estudo apresenta limitações, nomeadamente ao nível das habilitações literárias e nível de formação dos inquiridos, onde a grande maioria possui grau de mestre e estudos na área das Ciências, o que poderá enviesar os resultados.

Na avaliação do *chatbot* em ambiente controlado, as opções de resposta negativas “Às vezes” e “Nunca” registam, na generalidade, percentagens mais baixas, o que significa que nesta avaliação, o PALbot fornece mais respostas corretas do que erradas. Quanto ao Ciberdúvidas, de um modo geral, o insucesso aumenta comparativamente com o PALbot, uma vez que as percentagens para as mesmas opções de resposta, suprarreferidas, aumentam.

Na avaliação em ambiente livre a probabilidade de sucesso diminui em relação à avaliação suprarreferida, dado que a base de informação tanto do PALbot como do Ciberdúvidas ainda não contém um conhecimento suficientemente rico para funcionar com o mesmo grau de sucesso observado em ambiente controlado.

Em suma, a maioria dos indicadores estatísticos têm opções de respostas positivas com maior expressão no PALbot do que no Ciberdúvidas, em ambos os ambientes de avaliação. À exceção da variável “habilitações literárias” em ambiente livre, em que se regista uma maior expressão na opção “Sempre” no sucesso das respostas obtidas, para as diferentes habilitações literárias, no Ciberdúvidas.

De acordo com os resultados do inquérito, em termos de facilidade de acesso à informação através do PALbot, os inquiridos tendem a considerar eficaz a utilização desta ferramenta para obter respostas para as suas dúvidas. Foram elaboradas 383 perguntas, para as quais o PALbot obteve 251 respostas corretas, obtendo um valor de sucesso de 65,54%. De salientar que, aquando da avaliação de desempenho do PALbot, o primeiro candidato relevante recuperado obteve um valor de 85,17%.

Capítulo 5 – Conclusão

5.1. Principais conclusões

É impossível prever todas as dúvidas que os utilizadores poderão ter sobre a Língua Portuguesa, mas a partir dos dados do Ciberdúvidas é possível esclarecer a maioria. Este trabalho foca-se no desenvolvimento do PALbot usando a base de conhecimento do Ciberdúvidas para avaliar a eficácia do agente de conversação.

Em primeiro lugar, compararam-se os algoritmos do *framework* ALICE e do SSS como base de desenvolvimento do PALbot. Como o objetivo passa por obter a correspondência máxima de perguntas do utilizador de forma a devolver respostas corretas, o PALbot foi projetado com base no SSS que suporta a compreensão das palavras portuguesas enquanto que é necessário adicionar manualmente as palavras portuguesas na ALICE.

Para o PALbot usufruir da base de conhecimento da plataforma Ciberdúvidas com 25.896 conjuntos de perguntas e respostas relativamente a dúvidas de Língua Portuguesa, foi necessária a elaboração de métodos para processar os dados do conjunto de perguntas e respostas, de forma a integrá-las na base de conhecimento deste sistema de informação.

Na avaliação do PALbot onde o utilizador esclarece livremente qualquer tipo de dúvidas em relação à língua portuguesa, denota-se que o agente de conversação ainda não contém uma base de conhecimento suficientemente rica para funcionar com o mesmo grau de sucesso semelhante ao ambiente onde o utilizador elabora perguntas relacionadas com os tópicos fornecidos. Porém, 65,28% dos inquiridos concordam que o PALbot facilita quase sempre ou sempre o acesso à informação, ou seja, o mesmo melhora o esclarecimento de dúvidas sobre a Língua Portuguesa.

O que reforça a pertinência do trabalho é sobretudo os inquiridos concordarem que agentes de conversação como o PALbot facilitam o acesso à informação. A percentagem de respostas corretas de 65,54% revela que este *chatbot* cumpre com os objetivos do presente trabalho sendo, no entanto, ainda necessário trabalhar no desenvolvimento de bases de informação/conhecimento mais vastas e com capacidade de abrangerem, cada vez mais, um maior leque de respostas corretas a diferentes dúvidas.

5.2. Contributos para a comunidade científica e empresarial

O PALbot é um agente de conversação que foi baseado em *chatbots* que foram utilizados no Edgar e no Filipe. Sendo assim, com este trabalho demonstra-se que o SSS é versátil ao ponto de adquirir bons resultados tanto na Língua Portuguesa como na Língua Inglesa.

A procura de informação e a sua acessibilidade na plataforma Ciberdúvidas ficaria mais completa com a implementação do agente de conversação desenvolvido neste trabalho. Mostra-se também que o acesso à informação é realizado de forma mais eficaz do que com um sistema de procura tradicional.

5.3. Limitações do estudo

Uma das dificuldades encontradas na elaboração deste trabalho foi o acesso aos algoritmos. O estudo de comparação de algoritmos foi baseado na ALICE e no SSS. Apesar de se ter obtido resultados satisfatórios seria ideal que houvesse mais opções de escolha de forma a dar visão ao trabalho.

O inquérito ao ser respondido presencialmente, dificulta a dimensão da amostra, uma vez que é necessário testar o PALbot com solicitações de perguntas do inquirido. Esta ferramenta não foi disponibilizada em *online*, uma vez que existiam limitações de espaço no serviço *online*.

5.4. Propostas de investigação futura

Apesar do presente trabalho ter cumprido os objetivos, serão pertinentes investigações futuras passíveis de amplificar e potencializar os resultados obtidos, a saber:

- Processamento contínuo dos dados para as respostas serem generalizadas, uma vez as respostas que estão no corpus são respostas a perguntas do Ciberdúvidas;
- Aumentar ainda mais a base de conhecimento do PALbot;
- Desenvolver métodos de forma a lidar melhor com tópicos que não estejam na base conhecimento;
- Melhorar a funcionalidade da aprendizagem do algoritmo SSS de forma a adicionar as novas perguntas e respostas de forma transparente e com *runtime*.

Bibliografia

- Abdul-Kader, S. A., & Woods, J. (2017). Question answer system for online feedable new born Chatbot. *Intelligent Systems Conference (IntelliSys)*, 863-869. doi: 10.1109/IntelliSys.2017.8324231.
- Abran, A., Khelifi, A., Suryan, W., & Seffah, A. (2003). Consolidating the ISO usability models. *11th international software quality management conference*, 2003, 23-25.
Retrieved from: https://www.researchgate.net/publication/2850057_Consolidating_the_ISO_Usability_Models.
- Ameixa, D., Coheur, L., Fialho, P., & Quaresma, P. (2014, August). Luke, I am your father: dealing with out-of-domain requests by using movies subtitles. *International Conference on Intelligent Virtual Agents*, 13-21. doi:10.1007/978-3-319-09767-1_2.
- Benotti, L., Martínez, M. C., & Schapachnik, F. (2014). Engaging high school students using chatbots. *Proceedings of the 2014 conference on Innovation & technology in computer science education*, 63-68. doi: 10.1145/2591708.2591728
- Berkup, S. B. (2014). Working with generations X and Y in generation Z period: Management of different generations in business life. *Mediterranean Journal of Social Sciences*, 5(19), 218. doi: 10.5901/mjss.2014.v5n19p218.
- Bocklisch, T., Faulkner, J., Pawlowski, N., & Nichol, A. (2017). Rasa: Open-source language understanding and dialogue management. *arXiv preprint arXiv:1712.05181*.
Retrieved from <https://arxiv.org/abs/1712.05181>.
- Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22(5), 811-817. doi: 10.1017/S1351324916000243.
- dos Santos, D. R., Silva, F. S., & de Jesus Brito, A. (2016). Aspectos comportamentais do algoritmo GraphMaster para a implementação de chatterbots. *Anais do Congresso Acadêmico de Tecnologia e Informática (CATI)*, 2, 67-70.
Retrieved from https://www.researchgate.net/publication/329056550_Aspectos_comportamentais_do_algoritmo_GraphMaster_para_a_implementacao_de_chatterbots.
- dos Santos, L. F. A. P. (2015). *Chatbots: On Demand Creation of Conversational Agents*. Instituto Superior Técnico, Lisboa.
- Fialho, P., & Coheur, L. (2015, October). ChatWoz: chatting through a Wizard of Oz. *17th International ACM SIGACCESS Conference on Computers & Accessibility*, 423-424. doi: 10.1145/2700648.2811334.
- Fialho, P., Coheur, L., Curto, S., Cláudio, P., Costa, A., Abad, A. & Trancoso, I. (2013). Meet EDGAR, a tutoring agent at MONSERRATE. *51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 61-66.
Retrieved from <https://www.aclweb.org/anthology/P13-4011>.
- Fialho, P., Marques, R., Martins, B., Coheur, L. & Quaresma, P. (2016). INESC-ID@ ASSIN: Medição de Similaridade Semântica e Reconhecimento de Inferência Textual. *Linguística*, 8(2), 33-42. doi: 10.21814/lm.8.2.233.
- High, R. (2012). The era of cognitive systems: An inside look at IBM Watson and how it works. *IBM Corporation*.
Retrieved from: <http://www.redbooks.ibm.com/abstracts/redp4955.html?Open>

- Jung, B., & Kopp, S. (2003). Flurmax: An interactive virtual agent for entertaining visitors in a hallway. *International Workshop on Intelligent Virtual Agents*, 23-26. doi: 10.1017/S1351324916000243.
- Laureano, R. & Botelho, M. (2010). *SPSS-O Meu Manual de Consulta Rápida*. Lisboa: Edições Sílabo.
- Leuski, A., Patel, R. & Traum, D. (2006). Building effective question answering characters. *7th SIGdial Workshop on Discourse and Dialogue*, 18–27. doi: 10.3115/1654595.1654600.
- Marietto, M. D. G. B., de Aguiar, R. V., Barbosa, G. D. O., Botelho, W. T., Pimentel, E., França, R. D. S., & da Silva, V. L. (2013). Artificial intelligence markup language: a brief tutorial. *arXiv preprint arXiv:1307.3091*. Retrieved from <https://arxiv.org/abs/1307.3091v1>.
- Maroco, J., & Garcia-Marques, T. (2006). Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas?. *Laboratório de psicologia*, 65-90. Retrieved from <http://hdl.handle.net/10400.12/133>.
- McNeal, M. L., & Newyear, D. (2013). Chatbot creation options. *Library Technology Reports*, 49(8), 11-17. Retrieved from <https://journals.ala.org/index.php/ltr/article/view/4505/5283>.
- Mendes, A. & M., Coheur, L. (2013). When the Answer comes into Question: Survey and open issues. *Natural Language Engineering*, 19(1), 1-32. doi: 10.1017/S1351324911000350.
- Mendonça, V., Melo, F. S., Coheur, L., & Sardinha, A. (2017). A conversational agent powered by online learning. *16th Conference on Autonomous Agents and MultiAgent Systems*, 1637-1639. Retrieved from <https://dl.acm.org/purchase.cfm?id=3091388>.
- Nunes, M. & O'Neill, H. (2004). *Fundamental de UML*. 3ª Edição. Lisboa: FCA Editora.
- Pereira, M. J., Coheur, L., Fialho, P. & Ribeiro, R. (2016). Chatbots Greetings to Human-Computer Communication. *arXiv preprint arXiv:1609.06479*. Retrieved from <https://arxiv.org/abs/1609.06479>.
- Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *arXiv preprint arXiv:1704.04579*. Retrieved from <https://arxiv.org/abs/1704.04579>.
- Song, Y., Yan, R., Li, X., Zhao, D., & Zhang, M. (2016). Two are better than one: An ensemble of retrieval-and generation-based dialog systems. *arXiv preprint arXiv:1610.07149*. Retrieved from <https://arxiv.org/abs/1610.07149>.
- Vegesna, A., Jain, P., & Porwal, D. (2018). Ontology based Chatbot (For E-commerce Website). *International Journal of Computer Applications*, 179(4), 51-55. doi: 10.5120/ijca2018916215.
- Wallace, R. (2003). *The elements of AIML style*. São Francisco: A.L.I.C.E. AI Foundation.

Anexos

Anexo 1 – Questionário

Este questionário enquadra-se num trabalho de investigação no âmbito de uma dissertação de Mestrado em Gestão de Sistemas de Informação, realizada no ISCTE-IUL, que visa construir um *chatbot* que responde a perguntas sobre a língua portuguesa. O objetivo deste questionário é avaliar o *chatbot* e comparar o seu desempenho com uma interface de procura convencional. Os resultados obtidos serão utilizados apenas para fins académicos.

Obrigado pela sua colaboração!

1. Sexo

- Feminino
- Masculino

2. Idade

3. Habilitações Literárias

- 12.º ano
- Curso Tecnológico/Profissional/Outros (Nível III)
- Licenciatura
- Mestrado
- Doutoramento
- Outro

4. Indique a sua área de formação.

5. Já conversou alguma vez com um *chatbot*?

- Sim
- Não

Interação com o *chatbot* - avaliação controlada

Observe a lista que lhe é dada com potenciais tópicos. Procure obter informação sobre esses tópicos através de perguntas submetidas ao *chatbot* e através do site Ciberdúvidas

(<https://ciberduvidas.iscte-iul.pt/>).

6. Obteve a informação que procurava por parte do *chatbot*?

- Nunca
- Às vezes
- Geralmente
- Quase sempre
- Sempre

7. Obteve a informação que procurava por parte do Ciberdúvidas?

- Nunca
- Às vezes
- Geralmente
- Quase sempre
- Sempre

8. Dê a sua opinião

Interação com o *chatbot* - avaliação livre

Procure obter mais informação sobre a língua portuguesa através de perguntas submetidas ao *chatbot* e através do site Ciberdúvidas (<https://ciberduvidas.iscte-iul.pt/>).

9. Obteve a informação que procurava por parte do *chatbot*?

- Nunca
- Às vezes
- Geralmente
- Quase sempre
- Sempre

10. Obteve a informação que procurava por parte do Ciberdúvidas?

- Nunca
 Às vezes
 Geralmente
 Quase sempre
 Sempre

11. Dê a sua opinião

12. O *chatbot* facilita o acesso à informação.

- Nunca
 Às vezes
 Geralmente
 Quase sempre
 Sempre

13. Dê a sua opinião

Muito obrigado pela sua colaboração

Anexo 2 – Resultados da avaliação

Pergunta	IDCiberdúvidas da resposta	Sim	Não	Total
o que é um casal?	35121	1	0	1
á ou há?	34995	1	0	1
há ou à	34995	1	0	1
Ação ou Acção?	34953	0	4	4
bi-campeao ou bicampeao qual o corret?	34873	0	5	5
cabelereiro ou cabeleireiro?	34873	0	5	5
Pois, conclusivo ou explicativo?	34741	1	0	1
Promenor ou Pormenor?	34618	1	3	4

palavras proparoxítonas	34583	1	0	1
Qual é o sujeito e o predicativo do sujeito na frase “A personagem principal é a sementinha”?	34558	1	0	1
areeiro ou arieito?	34321	1	0	1
a cores ou a cores ou em cores?	34317	1	0	1
Deem ou dêem?	34311	1	3	4
Qual o plural de “cor de laranja”?	34096	0	5	5
um dia da caça e outro no caçador	33911	0	5	5
currículum ou currículo em língua portuguesa?	33910	1	0	1
hora ou ora?	33729	1	0	1
quai sao os pronomes de tratamento	33723	1	0	1
à prazo ou a prazo?	33413	1	0	1
a longo prazo ou em longo prazo?	33413	0	5	5
retificar ou ratificar têm o mesmo significado?	33403	1	0	1
retificar ou ratificar?	33403	1	0	1
o que é alteridade?	33300	1	0	1
Quando se diz obrigado e obrigada?	33198	0	5	5
presente ou futuro do conjuntivo	32967	1	0	1
O que significa abismar?	32764	0	5	5
Significado de coletivo	32721	1	0	1
tenão ou tensão	32693	1	0	1
o que é migração?	32693	0	5	5
O que é a natureza?	32570	1	0	1
substantivo coletivo da palavra habitantes	32570	1	2	3
apareceste ou apareces-te?	32532	0	5	5
o que são cores compostas?	32419	0	5	5
A palavra entreolhares pode ser um substantivo?	32325	0	5	5
acento ou assento?	32243	0	5	5
escreve-se Egito ou Egipto?	32129	1	0	1
Bem como ou assim como?	31767	0	5	5
Anexados / Em anexo	31723	1	0	1
o que significa anexo	31723	0	3	3
a vista ou à vista	31497	1	0	1
a nível de ou em nível de?	31447	1	0	1
supérgluo ou superfulo?	31447	0	5	5
Nubelado ou Nublado?	31429	1	0	1
mexer ou mecher?~	31338	1	0	1
Onde poderei aprender português?	31246	0	5	5
eu conto uma história ou estória?	31206	0	5	5
Na minha opinião pessoal ou na minha opinião?	31152	1	0	1
Vem ou Veem?	31059	1	0	1
fim-de-semana ou fim de semana?	31047	0	5	5
paralisação ou paralização?	31047	0	5	5
plural de pêra	31010	1	0	1

escreve-se por cento ou por cento?	30932	1	0	1
por cento ou por cento	30932	1	0	1
meteorologia ou metereologia	30842	1	1	2
significado de luso portugueses	30826	1	0	1
rúbrica	30783	1	0	1
diferença entre conselho e concelho	30584	0	4	4
diferença entre absolver ou absorver?	30584	0	5	5
escreve-se insisto que ou insisto em que	30555	1	0	1
núvem ou nuvem?	30496	1	0	1
Aparte e à parte?	30420	1	1	2
em que continente fica o camboja?	30298	0	5	5
matado ou morto?	30265	1	0	1
pode-se dizer eminente ou iminente?	30239	1	1	2
diferença entre haver ou a ver?	30195	0	5	5
Quais aos cursos que o ISCTE tem de licenciatura, mestrado e doutoramento?	30167	0	5	5
Enumeração dentro de um mesmo período	29970	0	5	5
como se escreve 600, seiscentos ou seissentos'	29849	0	4	4
o que é o que é Yoga?	29646	1	0	1
ziguezaguiar	29646	0	5	5
nao adianta chorar sobre leite derramado	29529	1	0	1
que tipo de ensino existe?	29388	0	5	5
Vir ou Vier?	29327	1	0	1
dupla grafia das palavras	29260	1	0	1
Diz-se no aguardp ou aguardo	29024	0	5	5
arrebetar ou rebentar?	28964	0	1	1
Escreve-se arrebentar ou rebentar?	28964	0	5	5
o que é um cão?	28755	0	5	5
qual é o antónimo de bondade?	28664	1	4	5
Qual é a diferença entre Antártida e àrtida?	28632	1	1	2
invertiu ou interveio?	28629	1	0	1
diz-me o que significa horripilo	28416	0	5	5
diz-se chichi ou xixi?	28411	0	5	5
escreve-se fição ou ficção?	28314	0	5	5
arrear ou arriar?	28302	1	0	1
Qual deles é o correto xérife ou xerife?	28058	1	0	1
qual a origem da palavra cruz?	28049	0	5	4
significado da palavra nicho	28027	1	0	1
conosco ou connosco	27752	1	0	1
qual está correto consertar ou concertar?	27491	1	0	1
um olho no peixe e outro no gato	27382	1	0	1
excreve-se à partida ou apartida?	26955	1	4	5
em quanto ou enquanto?	26843	1	0	1
repetir de novo ou só repetiur?	25876	0	4	4
fição ou ficação	25655	0	5	5
palarra membra existe?	25579	1	4	5
A palavra vário	25474	1	0	1

qual é o plural de caminhão-cisterna	25473	1	0	1
Caça à multa	25439	1	0	1
alcochete	25436	0	3	3
A expressão besta quadrada	25406	1	0	1
O conceito de escritório	25370	1	0	1
o que é um bisavô?	25073	1	4	5
Amanhã à de chover, escreve-se com à ou há?	25072	0	5	5
o que é o almoço amanhã?	25072	0	5	4
significado de sinfonia e biblioteca	25051	1	0	1
A sintaxe do verbo recordar	25041	1	0	1
pitula ou pilula	24100	0	1	1
paralisado ou paralizado?	23993	0	2	2
granito é que tipo de rocha?	23680	0	5	5
diferença entre parcela única ou de uma só vez	23354	0	5	5
obrigada ou obrigado?	23191	0	5	5
A nível de ou em níve de	23067	1	0	1
a nivel ou em nivel de?	23067	1	0	1
o que é um pc?	22969	1	0	1
significado de medeia?	22854	0	5	5
significado de visar?	22854	0	5	5
significado Ênclipse	22854	0	5	5
significado de desasnado	22854	0	5	5
significado de inócuo	22854	0	5	5
significado de xilema e floema	22854	0	5	5
quando uso lhe ou o	22524	0	5	5
E orações conclusivas?	22456	0	5	5
a palavra quites existe?	22058	0	5	5
palavra quite existe?	22058	0	5	5
diz-se chegar em ou chegar a ?	21904	1	0	1
mais vale um passaro na mao do que dois a voar	21375	1	0	1
deve haver ou devem haver?	20896	1	3	4
Qual deles o mais correto decerto ou de certo?	20889	1	0	1
boas festas ou boas-festas?	20112	1	0	1
vitor está	19339	0	5	5
penalizado e punido é o mesmo termo'	19304	0	5	5
há pouco ou a pouco'	19051	0	5	5
mecher ou mexer?	18915	0	2	2
de mais ou demais, posso usar para o mesmo significado?	18150	0	5	5
Vírgula entre sujeito e verbo	18070	1	0	1
açoriano ou açoreano?	17889	1	0	1
covadonga	17889	0	5	5
pachorra ou paxorra?	17712	0	1	1
Uso de aspas?	17255	1	1	2
o que é elemento quimico?	17255	0	5	5

qual é a diferença do nome próprio do comum?	17108	1	2	3
Organograma ou Organigrama?	17079	2	0	2
polissemia e homonímia	16862	1	0	1
definição de computador?	16850	1	0	1
puder ou poder?	16682	1	0	1
tempos verbais	16669	1	2	3
imigrante ou emigrante?	16450	0	5	4
diz-se coco ou côco?	16405	0	5	5
Propício ou propenso?	16404	0	5	5
o que é uma universidade?	16231	1	0	1
uso do verbo haver?	16106	1	1	2
tenho uma dúvida escreve-seacção ou acção?	16075	0	5	5
Deve-se escrever A ponto de ou ao ponto de?	16024	1	1	2
haja ou aja?	15956	1	4	5
O Que é insipido?	15949	0	5	5
o que é a internet?	15702	1	0	1
Devidas providências ou só providências?	15630	1	0	1
nebulosidade ou nebolusidade?	15619	1	0	1
beringela ou berinjala?	15376	1	0	1
adequa ou adequada	15308	1	0	1
o seguro morreu de velho	15306	0	5	5
significado mesóclise?	15036	0	5	5
para mim ou para eu?	14798	1	0	1
para eu ou para mim?	14798	1	0	1
Qual a diferença entre o português de Portugal e o do brasil?	14734	1	0	1
acerca de ou cerca de?	14640	1	0	1
o que é webmaster?	14582	2	0	2
senão ou se não	14518	1	0	1
significado de etc	14497	1	0	1
Uso em que ou onde?	14487	1	3	4
qual a forma correcta aonde ou onde?	14487	1	4	5
facto ou fato	14314	0	5	5
explique-me os modos verbais	14200	0	2	2
A champanhe/ o champanhe	14166	1	0	1
o qu quer dizer aproximar?	13321	0	4	4
em maos ou em mao?	12943	1	0	1
Posso dizer os dois, assistir o e assistir ao?	12876	1	0	1
femininp de primeiro ministro?	12832	0	5	5
Diz-se agradecer a ou agradecer por?	12786	1	0	1
agradecer pela ou a	12786	0	5	5
O ano escolar é Lectivo ou Letivo?	12719	1	0	1
has de ou has-de?	12703	1	0	1
o que significa enclise ou proclise?	12536	1	3	4
km ou qm	12467	1	0	1
como se escreve saem ou saiem?	12422	0	5	5

significado de agnóstico	12328	1	0	1
o que significa ciclo vicioso?	12160	1	0	1
Quando usar para mim ou para eu?	12110	0	5	5
quando usar onde e em que?	12110	0	5	5
qual é o significado de animal?	12107	1	3	4
de encontro ou ao encontro?	11999	1	0	1
ir de ou ao encontro'	11999	1	0	1
são suficientes ou é suficiente?	11940	0	5	5
é ré ou réu?	11782	0	4	4
O planeta é Úrano ou Urano?	11725	1	0	1
discriminar ou discriminar	11519	1	0	1
dispor e dispuser é igual?	11518	1	0	1
Encarar de frente ou só encarar?	11478	1	0	1
Como posso dizer, em duas metades iguais ou em duas partes parecidas?	11329	0	5	5
qual o correto saiem ou saem?	11065	0	5	5
o que é cucuco?	11065	0	5	5
o que é zina?	11065	0	5	5
e remocar?	11065	0	5	5
atossicar?	11065	0	5	5
eminente ou iminente?	10914	2	1	3
louvar o Senhor ou louvar ao Senhor?	10681	1	0	1
o que é uma análise critica?	10638	1	0	1
como se diz, comprimento ou cumprimento?	10559	1	0	1
o que é um sítio?	10491	1	0	1
diz se biliao ou mil milhoes?	10391	1	0	1
o que quer dizer Octeto	10264	1	4	5
sabes me dizer o que juvenelizante quer isto quer dizer	10264	0	5	5
e mendacioso?	10264	0	5	5
e odiento?	10264	0	5	5
a principio ou em principio?	10236	2	0	2
diz-se media ou medeia?	9893	1	1	2
indenmização ou indeminização?	9590	0	1	1
qual o mais correto, viagem ou viagem?	9350	0	5	5
pais natal ou pais natais?	9340	1	0	1
capelcapeloões ou capelães?	9340	0	5	5
Sociodemográfico ou socio-demográfico?	9244	1	0	1
mais grande ou maior?	8158	0	5	5
a par ou ao par?	8118	1	1	2
diz-se milhão ou milhoes?	8020	1	2	3
diz se jurí ou júri?	7931	0	5	5
separar sujeito e predicado	7742	1	0	1
Vírgula entre sujeito e predicado	7742	0	5	3
qual a definição de amigo?	7721	1	4	5
o que é um amigo?	7721	0	5	5
amigos amigos negocios a parte	7721	0	5	5
quais sao as regras de acentuação?	7652	0	5	5

encima ou em cima?	7336	1	0	1
qual é o sinónimo de envolvente?	6795	1	4	5
independente ou independetemente	6570	1	0	1
concelho ou conselho?	6568	0	4	4
Correr atrás do prejuízo	6428	1	0	1
infestado ou inpestado?	6187	0	2	2
utilizar ou utilizar	6160	1	0	1
qual a diferença, seção ou sessão	6095	1	0	1
escreve se o quanto antes ou quanto antes	6077	1	0	1
o que é gerúndio?	5857	1	2	3
uso do apóstrofo	5823	1	0	1
é mão-de-obra ou mão de obra?	5805	0	5	5
acentuação álcool e alcoólico	5596	1	0	1
quando usar obrigado ou obrigada?	5573	1	3	4
como conjugar a palavra implicar?	5376	0	5	5
horas extra ou extras?	5321	1	0	1
desapercebido ou despercebido	5257	1	0	1
qual é a correta despercebidos ou despercebidos?	5257	1	0	1
Desapercebido ou Despercebido?	5257	1	3	4
houveram ou houve	4905	1	1	2
houve ou houveram?	4905	2	2	4
usa-se estender ou extender?	4855	1	0	1
Laos é um país na Europa?	4757	0	5	5
Qual é a melhor maneira de uma pessoa expandir o seu vocabulário através de um estudo unicamente devotado a isso?	4721	1	0	1
o que é um holter?	4702	1	0	1
Receita sopa da pedra	4685	1	0	1
qual é o feminino de embaixador?	4679	1	0	1
como se escreve felipe ou filipe?	4670	1	0	1
camboja	4663	1	0	1
Qual é o antónimo de leveza?	4662	1	0	1
o que significa facial	4655	1	0	1
qual é o significado facial?	4655	1	0	1
significado facial	4655	0	5	5
Calibrar ou equilibrar	4562	1	0	1
Eu vou equilibrar o pneu ou calibrar?	4562	1	0	1
embaixo ou embaixo?	4557	1	1	2
casa-se ou casasse	4444	0	5	5
com certeza ou concerteza	4414	0	5	5
Escreve-se na medida em que ou à medida que	4353	1	0	1
junto a ou no ou ao	4280	1	0	1
à rua ou na rua?	4120	0	5	5
foda-se ou fodasse?	4066	0	1	1
mechilhão ou mexilhão?	4066	0	1	1
Há dois anos ou há dois anos atrás?	4038	1	0	1
escreve-se em vez de ou invés?	3979	1	4	5

bem ou bom?	3977	1	2	3
correcto ou correto?	3924	0	5	5
qual deste dois é o correto antárctida ou ártida?	3812	1	2	3
egipcio ou egicio	3697	0	5	4
precaer e prevenir é igual?	3594	1	4	5
diferença entre mal ou mau	3536	1	0	1
mal ou mau?	3536	1	0	1
Qual o mais correto preveram ou previram?	3433	1	4	5
é consiste de ou consiste em	3155	0	5	5
quiser ou quizer?	3153	1	0	1
é através de ou por meio de?	3082	1	1	2
chegar a casa ou chegar em casa?	3063	1	0	1
quimera?	2994	0	3	3
descrição ou discrição?	2882	1	0	1
Qual é a diferença entre descrição ou discrição	2882	1	0	1
viajem ou viagem?	2827	1	4	5
Posso usar multidão de pessoas?	2768	1	0	1
qual é o cletivo de avião?	2757	1	4	5
peço ou pesso?	2753	0	5	5
a fim ou afim?	2680	1	0	1
perperca e perda é ihgual	2638	0	5	5
significado de cujo	2606	1	0	1
responder a ou responder ao?	2594	1	0	1
que são palavras sinonimas?	2545	1	0	1
o que são Palavras aglutinadas	2545	0	5	5
o que significa transumanismo?	2447	1	0	1
qual é o feminino de politico?	2404	1	2	3
é online ou on-line?	2295	1	2	3
e lazeria?	2031	0	5	5
sisignificado de beneplácito	1957	0	1	1
quando uso perca ou perda?	1948	1	0	1
Qual a diferença entre palavras homógrafas e homófonas?	1903	1	0	1
o que são homónimos?	1903	1	4	5
paralavra voo	1762	1	0	1
vÉ vôo ou voo?	1762	0	5	5
Quando se deve usar um hiffen nas frase?	1721	0	5	5
diz-se islamita ou islamista?	1701	1	0	1
está certo usar desciminar e discriminar?	1661	1	1	2
de mais ou demais?	1659	1	0	1
haver e a ver?	1427	0	5	5
escreve-se quaduplicar ou quadruplicar?	1377	1	0	1
trás ou traz?	1327	0	5	5
onde fica sidney	1293	1	3	4
hera ou era?	1293	0	5	5
Escreve-se a grama ou o grama?	1193	1	1	2

meio-dia e meio ou meio-dia e meia?	1102	1	0	1
Comparação ente por que ou porque	981	0	5	5
O que é o acordo ortográfico	969	1	0	1
explique me entreviu e interveio	911	1	1	2
excreve-se meio feita ou meia feita?	879	1	0	1
qual o amis correto apesar de o ou apesar do?	868	1	0	1
coser ou cozer?	750	1	0	1
O de captura é Mandado ou Mandato?	694	1	0	1
mandado ou mandato?	694	2	0	2
demais ou de mais?	608	1	0	1
A domicilio ou em domicilio?	594	1	0	1
Qual é a forma correta de dizer a domicilio ou em domicilio?	594	1	0	1
Aonde existe?	548	1	0	1
onde ou aonde?	548	1	0	1
aonde ou onde	548	0	5	5
NATO ou OTAN é a mesma instituição?	520	1	0	1
a gente ou agente?	494	1	1	2
o que significa slb?	462	0	5	5
O vocábulo facto mantém o c?	461	1	0	1
aceitado, aceite e aceito	441	1	0	1
Aceita-se ou aceitam-se?	441	0	5	5
tem a ver ou tem que ver?	253	1	0	1
Porquê e por quê	243	1	3	4
bem vindo ou bem-vindo	211	1	0	1
O uso do hífen	112	1	0	1
Quando se usa suspeito ou suspeita?	86	1	0	1
corrimãos ou corrimões?	70	1	0	1
A gente ou agente?	0	1	0	1
A principio ou em principio?	0	1	0	1
aja e haja?	0	1	0	1
ao encontro de ou ao de encontro a?	0	1	0	1
comprimento ou cumprimento?	0	1	0	1
descrição ou discrição	0	1	0	1
Descrição ou discrição?	0	1	0	1
Deve se escrever Afim ou a fim?	0	1	0	1
diz se descriminar ou discriminar?	0	1	0	1
Diz-se a champnhe ou o champanhe?	0	1	0	1
Diz-se a cores ou em cores?	0	1	0	1
é viagem ou viagem?	0	1	0	1
Escreve-se a domicilio ou em domicilio?	0	1	0	1
hora e ora são iguais?	0	1	0	1
Na medida e que ou à medida que?	0	1	0	1
o que é a sopa da pedra?	0	1	0	1
Quanto ao responder posso dizer responder o ou ao responder ao?	0	1	0	1
Senão ou se nao?	0	1	0	1
Ao encontro de ou de encontro a ?	0	1	0	1

atravé ou por meio pode-se usar da mesma forma?	0	1	0	1
por hora ou por ora?	0	1	0	1
A prazo à prazo?	0	1	0	1
vem ou veem?	0	1	0	1
responder o ou ao	0	1	0	1
diz-se em baixo ou embaixo?	0	1	0	1
no brasil é mais comum filipe ou felipe?	0	1	1	2
quando uso senao ou se nao?	0	1	1	2
quando se usa dar ou dá?	0	1	4	5
quando uso hifens?	0	1	4	5
qual deles o correto segmento ou seguimento?	0	1	4	5
forma correta a meu ver ou ao meu ver?	0	0	5	5
Qual o significado de gótico?	0	0	5	5
ao meu ver ou a meu ver?	0	0	5	5
qual é o acordo ortográfico mais atual?	0	0	5	5

Anexo 3 – Caracterização da pesquisa de perguntas no Cibérduvidas que o PALbot que o PALbot forneceu resposta correta

Pergunta	Ciberdúvidas	
	Resposta correta	Hiperligação da pesquisa
escreve-se em vez ou invés	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=em+vez+ou+inv%C3%A9s
escreve-se à partida ou apartida?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=%C3%A0+partida+ou+apartida
haja ou aja?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=haja+ou+aja
o que é um bisavô?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=bisav%C3%B4
o que são homónimos?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=hom%C3%B3nimos
palavra membra existe?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=palavra+membra+
precaver e prevenir é igual?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=precaver+e+prevenir+
qual a definição de amigo?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=defini%C3%A7%C3%A3o+de+amigo
qual a forma correcta aonde ou onde?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=forma+correcta+aonde+ou+onde

qual é o antónimo de bondade?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=ant%C3%B3nimo+de+bondade
qual é o sinónimo de envolvente?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=sin%C3%B3nimo+de+envolvente
Qual o mais correto preveram ou previram?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=previram+ou+previram
quando se usa dar ou dá?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=dar+ou+d%C3%A1
quando uso hifens?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=hifens
viagem ou viagem?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=viagem+ou+viagem
qual deles o correto segmento ou seguimento?	Não	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=segmento+ou+seguimento
o que quer dizer Octeto	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=octeto
qual é o coletivo de avião?	Sim	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=coletivo+de+aviao

Anexo 4 – Caracterização da pesquisa de perguntas no Cibéduvidas que PALbot que o PALbot não forneceu resposta correta

Pergunta	Ciberdúvidas	
	Hiperligação da pesquisa	Resposta correta
diz se jurí ou júri?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=jur%C3%AD+ou+j%C3%BAri	Não
casa-se ou casasse	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=casa-se+ou+casasse	Não
haver e a ver?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=haver+e+a+ver	Não
mais grande ou maior?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=mais+grande+ou+maior	Sim
o que significa slb?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=slb	Não
à rua ou na rua?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=%C3%A0+rua+ou+na+rua	Não
agradecer pela ou a	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=agradecer+pela+ou+a	Não
Propício ou propenso?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=Prop%C3%ADcio+ou+propenso	Não

paralisação ou paralização?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=paralisa%C3%A7%C3%A3o+ou+paraliza%C3%A7%C3%A3o	Não
O que é um cão?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=c%C3%A3o	Não
a palavra quites existe?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=quites	Não
escreve-se fição ou ficção?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=fi%C3%A7%C3%A3o+ou+fic%C3%A7%C3%A3o	Sim
significado de inócuo	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=significado+de+in%C3%B3cuo	Não
significado de desasnado	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=desasnado	Não
Vírgula entre sujeito e predicado	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=V%C3%ADrgula+entre+sujeito+e+predicado	Sim
significado facial	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=significado+facal	Não
Quando se deve usar um hiffen nas frases?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=hiffen+nas+frases	Não
ao meu ver ou a meu ver?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=ao+meu+ver+ou+a+meu+ver	Não
supérgluo ou superfulo?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=sup%C3%A9rgluo+ou+superfulo	Não
significado de medeia?	https://ciberduvidas.iscte-iul.pt/pesquisa/consultorio?termo=significado+de+medeia	Não

Anexo 5 – Caracterização genérica da procura de informação pelo PALbot e pelo Ciberdúvidas em ambiente controlado

		Contagem	% de N da coluna da camada
Obteve a informação que procurava por parte do chatbot?	Nunca	3	4,10%
	Às vezes	5	6,80%
	Geralmente	25	34,20%
	Quase sempre	27	37,00%
	Sempre	13	17,80%
	Total	73	100,00%
Obteve a informação que procurava por parte do Ciberdúvidas?	Nunca	4	5,50%
	Às vezes	15	20,50%
	Geralmente	22	30,10%
	Quase sempre	26	35,60%
	Sempre	6	8,20%
	Total	73	100,00%

Anexo 6 – Caracterização genérica da procura de informação pelo PALbot e pelo Ciberdúvidas em ambiente livre

		Contagem	% de N da coluna da camada
		Obteve a informação que procurava por parte do chatbot?	Nunca
Às vezes	31		42,50%
Geralmente	22		30,10%
Quase sempre	7		9,60%
Sempre	6		8,20%
Total	73		100,00%
Obteve a informação que procurava por parte do Ciberdúvidas?	Nunca	17	23,03%
	Às vezes	25	34,20%
	Geralmente	21	28,80%
	Quase sempre	7	9,60%
	Sempre	3	4,10%
	Total	73	100,0%

Anexo 7 – Relação entre idade e a avaliação em ambiente controlado do PALbot

		Obteve a informação que procurava por parte do <i>chatbot</i> ?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Idade	Geração X	Frequência de resposta	0	1	2	3	1	7
		Percentagem de resposta	0,00%	20,00%	8,00%	11,54%	7,69%	9,72%
	Geração Y	Frequência de resposta	2	4	16	18	9	49
		Percentagem de resposta	66,67%	80,00%	64,00%	69,23%	69,23%	68,06%
	Geração Z	Frequência de resposta	1	0	7	5	3	16
		Percentagem de resposta	33,33%	0,00%	28,00%	19,23%	23,08%	22,22%
	Total	Frequência de resposta	3	5	25	26	13	72
		Percentagem de resposta	4,17%	6,94%	34,72%	36,11%	18,06%	100,00%

Anexo 8 – Relação entre idade e a avaliação em ambiente controlado do Ciberdúvidas

		Obteve a informação que procurava por parte do Ciberdúvidas?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Idade	Geração X	Frequência de resposta	0	3	1	2	1	7
		Percentagem de resposta	0,00%	20,00%	4,55%	8,00%	16,67%	9,72%
	Geração Y	Frequência de resposta	3	9	15	18	4	49
		Percentagem de resposta	75,00%	60,00%	68,18%	72,00%	66,67%	68,06%
	Geração Z	Frequência de resposta	1	3	6	5	1	16
		Percentagem de resposta	25,00%	20,00%	27,27%	20,00%	16,67%	22,22%
Total	Frequência de resposta	4	15	22	25	6	72	
	Percentagem de resposta	5,56%	20,83%	30,56%	34,72%	8,33%	100,00%	

Anexo 9 – Relação entre habilitações literárias e a avaliação em ambiente controlado do PALbot

		Obteve a informação que procurava por parte do chatbot?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Habilitações Literárias	12.º ano	Frequência de resposta	1	1	6	10	4	22
		Percentagem de resposta	1,4%	1,4%	8,2%	13,7%	5,5%	30,1%
	Curso Tecnológico/Profissional/Outros (Nível III)	Frequência de resposta	0	0	0	0	1	1
		Percentagem de resposta	0,0%	0,0%	0,0%	0,0%	1,4%	1,4%
	Licenciatura	Frequência de resposta	2	0	7	7	4	20
		Percentagem de resposta	2,7%	0,0%	9,6%	9,6%	5,5%	27,4%
	Mestrado	Frequência de resposta	0	4	10	8	3	25
		Percentagem de resposta	0,0%	5,5%	13,7%	11,0%	4,1%	34,2%
	Doutoramento	Frequência de resposta	0	0	2	2	1	5

Total	Percentagem de resposta	0,0%	0,0%	2,7%	2,7%	1,4%	6,8%
	Frequência de resposta	3	5	25	27	13	73
	Percentagem de resposta	4,1%	6,8%	34,2%	37,0%	17,8%	100,0%

Anexo 10 – Relação entre habilitações literárias e a avaliação em ambiente controlado do Ciberdúvidas

		Obteve a informação que procurava por parte do Ciberdúvidas?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Habilitações Literárias	12.º ano	Frequência de resposta	2	4	8	8	0	22
		Percentagem de resposta	2,7%	5,5%	11,0%	11,0%	0,0%	30,1%
	Curso Tecnológico/Profissional/Outros (Nível III)	Frequência de resposta	0	0	0	0	1	1
		Percentagem de resposta	0,0%	0,0%	0,0%	0,0%	1,4%	1,4%
	Licenciatura	Frequência de resposta	2	2	6	8	2	20
		Percentagem de resposta	2,7%	2,7%	8,2%	11,0%	2,7%	27,4%
	Mestrado	Frequência de resposta	0	7	7	9	2	25
		Percentagem de resposta	0,0%	9,6%	9,6%	12,3%	2,7%	34,2%
	Doutoramento	Frequência de resposta	0	2	1	1	1	5
		Percentagem de resposta	0,0%	2,7%	1,4%	1,4%	1,4%	6,8%
	Total	Frequência de resposta	4	15	22	26	6	73
		Percentagem de resposta	5,5%	20,5%	30,1%	35,6%	8,2%	100,0%

Anexo 11 – Relação entre habilitações literárias e a avaliação em ambiente livre do PALbot

		Obteve a informação que procurava por parte do chatbot?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Habilitações Literárias	12.º ano	Frequência de resposta	3	7	8	2	2	22
		Percentagem de resposta	4,1%	9,6%	11,0%	2,7%	2,7%	30,1%
	Curso Tecnológico/Profissional/Outros (Nível III)	Frequência de resposta	0	0	0	1	0	1
		Percentagem de resposta	0,0%	0,0%	0,0%	1,4%	0,0%	1,4%
	Licenciatura	Frequência de resposta	4	8	6	1	1	20
		Percentagem de resposta	5,5%	11,0%	8,2%	1,4%	1,4%	27,4%
	Mestrado	Frequência de resposta	0	14	7	2	2	25
		Percentagem de resposta	0,0%	19,2%	9,6%	2,7%	2,7%	34,2%
	Doutoramento	Frequência de resposta	0	2	1	1	1	5
		Percentagem de resposta	0,0%	2,7%	1,4%	1,4%	1,4%	6,8%
	Total	Frequência de resposta	7	31	22	7	6	73
		Percentagem de resposta	9,6%	42,5%	30,1%	9,6%	8,2%	100,0%

Anexo 12 – Relação entre habilitações literárias e a avaliação em ambiente livre do Ciberdúvidas

		Obteve a informação que procurava por parte do Ciberdúvidas?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Habilitações Literárias	12.º ano	Frequência de resposta	8	5	8	1	0	22
		Percentagem de resposta	11,0%	6,8%	11,0%	1,4%	0,0%	30,1%
	Curso Tecnológico/Profissional/Outros (Nível III)	Frequência de resposta	0	0	0	1	0	1
		Percentagem de resposta	0,0%	0,0%	0,0%	1,4%	0,0%	1,4%
	Licenciatura	Frequência de resposta	5	8	5	2	0	20
		Percentagem de resposta	6,8%	11,0%	6,8%	2,7%	0,0%	27,4%
	Mestrado	Frequência de resposta	3	12	7	1	2	25
		Percentagem de resposta	4,1%	16,4%	9,6%	1,4%	2,7%	34,2%

Doutoramento	Frequência de resposta	1	0	1	2	1	5
	Percentagem de resposta	1,4%	0,0%	1,4%	2,7%	1,4%	6,8%
Total	Frequência de resposta	17	25	21	7	3	73
	Percentagem de resposta	23,3%	34,2%	28,8%	9,6%	4,1%	100,0%

Anexo 13 – Relação entre áreas de formação e avaliação em ambiente controlado do PALbot

		Obteve a informação que procurava por parte do chatbot?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Indique a sua área de formação	Artes	Frequência de resposta	1	0	1	3	0	5
		Percentagem de resposta	1,4%	0,0%	1,4%	4,1%	0,0%	6,8%
	Ciências	Frequência de resposta	0	2	7	9	8	26
		Percentagem de resposta	0,0%	2,7%	9,6%	12,3%	11,0%	35,6%
	Economia	Frequência de resposta	0	0	6	2	0	8
		Percentagem de resposta	0,0%	0,0%	8,2%	2,7%	0,0%	11,0%
	Engenharias	Frequência de resposta	1	1	8	5	3	18
		Percentagem de resposta	1,4%	1,4%	11,0%	6,8%	4,1%	24,7%
	Humanidades	Frequência de resposta	1	2	3	8	2	16
		Percentagem de resposta	1,4%	2,7%	4,1%	11,0%	2,7%	21,9%
	Total	Frequência de resposta	3	5	25	27	13	73
		Percentagem de resposta	4,1%	6,8%	34,2%	37,0%	17,8%	100,0%

Anexo 14 – Relação entre áreas de formação e avaliação em ambiente controlado do Ciberdúvidas

		Obteve a informação que procurava por parte do Ciberdúvidas?					Total	
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre		
Indique a sua área de formação	Artes	Frequência de resposta	1	0	0	4	0	5
		Percentagem de resposta	1,4%	0,0%	0,0%	5,5%	0,0%	6,8%
	Ciências	Frequência de resposta	0	5	9	8	4	26
		Percentagem de resposta	0,0%	6,8%	12,3%	11,0%	5,5%	35,6%
	Economia	Frequência de resposta	0	1	4	3	0	8
		Percentagem de resposta	0,0%	1,4%	5,5%	4,1%	0,0%	11,0%
	Engenharias	Frequência de resposta	1	4	5	6	2	18
		Percentagem de resposta	1,4%	5,5%	6,8%	8,2%	2,7%	24,7%
	Humanidades	Frequência de resposta	2	5	4	5	0	16
		Percentagem de resposta	2,7%	6,8%	5,5%	6,8%	0,0%	21,9%
	Total	Frequência de resposta	4	15	22	26	6	73
		Percentagem de resposta	5,5%	20,5%	30,1%	35,6%	8,2%	100,0%

Anexo 15 – Relação entre áreas de formação e avaliação em ambiente livre do PALbot

		Obteve a informação que procurava por parte do chatbot?					Total	
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre		
Indique a sua área de formação	Artes	Frequência de resposta	1	2	2	0	0	5
		Percentagem de resposta	1,4%	2,7%	2,7%	0,0%	0,0%	6,8%
	Ciências	Frequência de resposta	1	10	11	2	2	26
		Percentagem de resposta	1,4%	13,7%	15,1%	2,7%	2,7%	35,6%
	Economia	Frequência de resposta	1	5	2	0	0	8
		Percentagem de resposta	1,4%	6,8%	2,7%	0,0%	0,0%	11,0%
	Engenharias	Frequência de resposta	2	7	3	3	3	18
		Percentagem de resposta	2,7%	9,6%	4,1%	4,1%	4,1%	24,7%
	Humanidades	Frequência de resposta	2	7	4	2	1	16

Total	Percentagem de resposta	2,7%	9,6%	5,5%	2,7%	1,4%	21,9%
	Frequência de resposta	7	31	22	7	6	73
	Percentagem de resposta	9,6%	42,5%	30,1%	9,6%	8,2%	100,0%

Anexo 16 – Relação entre áreas de formação e avaliação em ambiente livre do Ciberdúvidas

		Obteve a informação que procurava por parte do Ciberdúvidas?					Total
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	
Artes	Frequência de resposta	1	2	1	1	0	5
	Percentagem de resposta	1,4%	2,7%	1,4%	1,4%	0,0%	6,8%
Ciências	Frequência de resposta	2	10	11	2	1	26
	Percentagem de resposta	2,7%	13,7%	15,1%	2,7%	1,4%	35,6%
Economia	Frequência de resposta	3	3	1	1	0	8
	Percentagem de resposta	4,1%	4,1%	1,4%	1,4%	0,0%	11,0%
Engenharias	Frequência de resposta	3	6	5	2	2	18
	Percentagem de resposta	4,1%	8,2%	6,8%	2,7%	2,7%	24,7%
Humanidades	Frequência de resposta	8	4	3	1	0	16
	Percentagem de resposta	11,0%	5,5%	4,1%	1,4%	0,0%	21,9%
Indique a sua área de formação	Frequência de resposta	17	25	21	7	3	73
	Percentagem de resposta	23,3%	34,2%	28,8%	9,6%	4,1%	100,0%

Anexo 17 – Avaliação genérica em ambiente livre do PALbot e do Ciberdúvidas

		Contagem	% de N da coluna da camada
		Obteve a informação que procurava por parte do chatbot?	Nunca
	Às vezes	31	42,5%
	Geralmente	22	30,1%
	Quase sempre	7	9,6%
	Sempre	6	8,2%
	Total	73	100,0%
Obteve a informação que procurava por parte do Ciberdúvidas?	Nunca	17	23,3%
	Às vezes	25	34,2%
	Geralmente	21	28,8%
	Quase sempre	7	9,6%
	Sempre	3	4,1%
	Total	73	100,0%

Anexo 18 – Relação entre idade e a avaliação em ambiente livre do PALbot

		Obteve a informação que procurava por parte do <i>chatbot</i> ?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Idade	Geração X	Frequência de resposta	0	4	2	0	1	7
		Percentagem de resposta	0,00%	12,90%	9,52%	0,00%	16,67%	9,72%
	Geração Y	Frequência de resposta	7	22	13	3	4	49
		Percentagem de resposta	100,00%	70,97%	61,90%	42,86%	66,67%	68,06%
	Geração Z	Frequência de resposta	0	5	6	4	1	16
		Percentagem de resposta	0,00%	16,13%	28,57%	57,14%	16,67%	22,22%
	Total	Frequência de resposta	7	31	21	7	6	72
		Percentagem de resposta	9,72%	43,06%	29,17%	9,72%	8,33%	100,00%

Anexo 19 – Relação entre idade e a avaliação em ambiente livre do Ciberdúvidas

		Obteve a informação que procurava por parte do Ciberdúvidas?						
		Nunca	Às vezes	Geralmente	Quase sempre	Sempre	Total	
Idade	Geração X	Frequência de resposta	1	2	2	1	1	7
		Percentagem de resposta	5,88%	8,00%	9,52%	16,67%	33,33%	9,72%
	Geração Y	Frequência de resposta	12	21	12	2	2	49
		Percentagem de resposta	70,59%	84,00%	57,14%	33,33%	66,67%	68,06%
	Geração Z	Frequência de resposta	4	2	7	3	0	16
		Percentagem de resposta	23,53%	8,00%	33,33%	50,00%	0,00%	22,22%
	Total	Frequência de resposta	17	25	21	6	3	72
		Percentagem de resposta	23,61%	34,72%	29,17%	8,33%	4,17%	100,00%

Anexo 20 – Relação entre idade com o histórico de interações com o PALbot

		Já conversou alguma vez com um <i>chatbot</i> ?			
		Não	Sim	Total	
Idade	Geração X	Frequência de resposta	2	5	7
		Percentagem de resposta	6,06%	12,82%	9,72%
	Geração Y	Frequência de resposta	21	28	49
		Percentagem de resposta	63,64%	71,79%	68,06%
	Geração Z	Frequência de resposta	10	6	16
		Percentagem de resposta	30,30%	15,38%	22,22%
Total	Frequência de resposta	33	39	72	
	Percentagem de resposta	45,83%	54,17%	100,00%	

Anexo 21 – Relação entre idade e a acessibilidade de informação fornecida pelo PALbot

		O chatbot facilita o acesso à informação.				Total	
		Às vezes	Geralmente	Quase sempre	Sempre		
Idade	Geração X	Frequência de resposta	1	2	4	0	7
		Percentagem de resposta	11,11%	12,50%	11,11%	0,00%	9,72%
	Geração Y	Frequência de resposta	7	12	22	8	49
		Percentagem de resposta	77,78%	75,00%	61,11%	72,73%	68,06%
	Geração Z	Frequência de resposta	1	2	10	3	16
		Percentagem de resposta	11,11%	12,50%	27,78%	27,27%	22,22%
Total	Frequência de resposta	9	16	36	11	72	
	Percentagem de resposta	12,50%	22,22%	50,00%	15,28%	100,00%	

Anexo 22 – Relação entre o histórico de iterações com a facilidade do acesso à informação fornecida pelo PALbot

		Já conversou alguma vez com um <i>chatbot</i> ?			
		Não	Sim	Total	
O <i>chatbot</i> facilita o acesso à informação.	Às vezes	Frequência de resposta	4	5	9
		Percentagem de resposta	5,48%	6,85%	12,33%
	Geralmente	Frequência de resposta	8	8	16
		Percentagem de resposta	10,96%	10,96%	21,92%
	Quase sempre	Frequência de resposta	18	19	37
		Percentagem de resposta	24,66%	26,03%	50,68%
	Sempre	Frequência de resposta	3	8	11
		Percentagem de resposta	4,11%	10,96%	15,07%
	Total	Frequência de resposta	33	40	73
		Percentagem de resposta	45,21%	54,79%	100,00%