



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CAMPUS ARARANGUÁ OU CENTRO DE CIÊNCIAS, TECNOLOGIA E SAÚDE  
CURSO DE GRADUAÇÃO EM ENGENHARIA DA COMPUTAÇÃO

Matheus Letzov Peloso

**Utilizando tecnologias *chatbot* para responder dúvidas sobre as vacinas do  
COVID-19**

Araranguá  
2022

Matheus Letzov Peloso

**Utilizando tecnologias *chatbot* para responder dúvidas sobre as vacinas do  
COVID-19**

Trabalho de Conclusão de Curso do Curso de Graduação em Engenharia da Computação do Campus Araranguá ou Centro de Ciências, Tecnologia e Saúde da Universidade Federal de Santa Catarina para a obtenção do título de Bacharel em Engenharia da Computação.

Orientador: Prof. Alison R. Panisson, Dr.

Araranguá

2022

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Letzov, Matheus

Utilizando tecnologias chatbot para responder dúvidas  
sobre as vacinas do COVID-19 / Matheus Letzov ;  
orientador, Alison Roberto Panisson, 2022.

56 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Campus Araranguá,  
Graduação em Engenharia de Computação, Araranguá, 2022.

Inclui referências.

1. Engenharia de Computação. 2. Tecnologias Chatbot. 3.  
Vacinas Covid-19. I. Panisson, Alison Roberto. II.  
Universidade Federal de Santa Catarina. Graduação em  
Engenharia de Computação. III. Título.

Matheus Letzov Peloso

**Utilizando tecnologias *chatbot* para responder dúvidas sobre as vacinas do  
COVID-19**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharel em Engenharia da Computação” e aprovado em sua forma final pelo Curso de Graduação em Engenharia da Computação.

Araranguá, 10 de Março de 2022.

---

Prof. Analucia Schiaffino Morales, Dra.  
Coordenador do Curso

**Banca Examinadora:**

---

Prof. Alison R. Panisson, Dr.  
Orientador

---

Prof. Anderson Luiz Fernandes Perez, Dr.  
Avaliador  
Instituição UFSC

---

Prof. Marcelo Daniel Berejuck, Dr.  
Avaliador  
Instituição UFSC

Este trabalho é dedicado aos meus colegas de classe e aos meus queridos pais.

## **AGRADECIMENTOS**

Primeiramente, gostaria de dedicar e agradecer aos meus pais que me apoiaram e acreditaram em mim durante toda essa longa jornada dentro da Universidade Federal de Santa Catarina. A trajetória foi cheia de adversidades e imprevistos, mas sem o apoio dos meus pais tudo seria muito mais difícil do que foi. Também gostaria de agradecer aos professores que me incentivaram e ajudaram durante todo o período da graduação, seja tirando uma dúvida ao final da aula ou ajudando com algum trabalho da disciplina. Gostaria também de agradecer ao meu orientador, por ter me auxiliado integralmente durante todo o desenvolvimento do TCC. Por fim, vale um caloroso agradecimento a todos os meus amigos que estiveram ao meu lado de maneira fiel e íntegra durante todo o tempo que estudei na UFSC. Não teria chegado tão longe sem eles. Ana, Calvin, Eriel, Jhonathan e Luiz em especial, me ajudaram em diversos momentos, sem eles a estrada teria chegado a um fim prematuro. Eternamente grato por tudo que fizeram para que a minha jornada na UFSC fosse mais a mais agradável e inesquecível possível. Junto daqueles que me apoiaram cheguei onde cheguei. Uma experiência única e inesquecível chega ao fim.

## RESUMO

Chatbots são softwares capazes de interagir com o usuário e compreender mensagens em linguagem natural. São várias as vantagens do uso de tecnologias chatbots. Uma das vantagens é a capacidade de interação dessas tecnologias, o que pode ser utilizado para sanar dúvidas e questionamentos de usuários sobre um determinado assunto. No contexto da pandemia da covid-19, onde existem diversas dúvidas acerca das vacinas e suas respectivas características, uma ferramenta como um chatbot pode ser útil e eficaz em fornecer informações. Com esse intuito, esse trabalho descreve uma investigação sobre tecnologias chatbots, e um estudo de caso sobre o desenvolvimento de um chatbot voltado a responder dúvidas acerca das vacinas da covid-19. Em nosso estudo de caso, realizamos diversas etapas, iniciando com a coleta das principais dúvidas que as pessoas possuem sobre as vacinas da covid-19, o levantamento de informações necessárias para construir a base de conhecimento do chatbot para responder as dúvidas levantadas, o desenvolvimento do chatbot, e um teste empírico do mesmo. Nosso estudo demonstra que tecnologias chatbots podem ser utilizadas de forma eficiente para fornecer informações de forma interativa sobre um tema atual e relevante.

**Palavras-chave:** Inteligência Artificial. Tecnologias Chatbots. Vacinas Covid-19.

## **ABSTRACT**

Chatbots are technologies capable of interacting with users in natural language. There are several advantages of using chatbot technologies. One of the advantages is the ability of these technologies to engage in dialogues with users, which can be used to provide answers to questions from users about a particular subject. In the context of the covid-19 pandemic, where there are several doubts about vaccines and their respective characteristics, a tool such as a chatbot can be useful and effective in providing information. To that end, this work describes an investigation about chatbot technologies, and a case study on the development of a chatbot aimed at answering questions about covid-19 vaccines. In our case study, we carried out several steps, starting with collecting the main questions that people have about covid-19 vaccines, collecting the information necessary to build the chatbot's knowledge base to answer those questions, developing of the chatbot, and an empirical evaluation of it. Our study demonstrates that chatbot technologies can be used efficiently to provide information interactively on a current and relevant topic.

**Keywords:** Artificial Intelligence. Chatbot Technologies. COVID-19 Vaccines.

## LISTA DE FIGURAS

Figura 1 – Word2Vec . . . . .	22
Figura 2 – Exemplos de entrada na NLU. . . . .	26
Figura 3 – Exemplo de interface de desenvolvimento do Rasa . . . . .	28
Figura 4 – Exemplo de treinamento NLU para perguntas coletadas. . . . .	32
Figura 5 – Exemplo treinamento NLU para comunicação amigável. . . . .	33
Figura 6 – Story sobre os efeitos colaterais de uma vacina específica. . . . .	39
Figura 7 – <i>Story</i> sobre os efeitos colaterais de todas as vacinas. . . . .	40
Figura 8 – Implementação da validação de nomes de vacinas. . . . .	41
Figura 9 – Implementação de checkpoints. . . . .	42
Figura 10 – Storie sobre a eficácia da primeira dose . . . . .	42
Figura 11 – Stories sobre a eficácia da segunda dose. . . . .	43
Figura 12 – Stories sobre o intervalo entre as doses das vacinas. . . . .	43
Figura 13 – Stories sobre o intervalo entre as doses das vacinas imediato. . . . .	44
Figura 14 – Stories sobre eficácia das vacinas por faixa etária. . . . .	44
Figura 15 – Stories sobre eficácia das vacinas por faixa etária imediata. . . . .	45
Figura 16 – Stories eficácia contra variantes . . . . .	45
Figura 17 – Stories Tecnologia vacinas . . . . .	46
Figura 18 – Stories Tecnologia vacinas imediata . . . . .	46
Figura 19 – Matriz de confusão com resultados dos testes da NLU. . . . .	49
Figura 20 – Histograma de distribuição de confiança no reconhecimento de intenções. . . . .	50
Figura 21 – Exemplo de <i>story</i> de teste desenvolvida. . . . .	51
Figura 22 – Matriz de confusão com resultado dos testes de stories. . . . .	52

## LISTA DE TABELAS

Tabela 1 – Principais dúvidas levantadas a partir de questionários. . . . .	31
Tabela 2 – Tabela de tecnologia usada pelas vacinas da COVID-19. . . . .	35
Tabela 3 – Tabela de Eficácia das Vacinas contra COVID-19. . . . .	36
Tabela 4 – Tabela de Intervalo entre as doses das Vacinas contra COVID-19. . . .	36
Tabela 5 – Tabela de Efeitos colaterais das Vacinas contra COVID-19. . . . .	37
Tabela 6 – Tabela de Eficácia das Vacinas contra COVID-19 por faixa etária. . . .	38
Tabela 7 – Tabela de Eficácia das Vacinas contra variantes da COVID-19. . . . .	38

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
1.1	OBJETIVOS	12
<b>1.1.1</b>	<b>Objetivo Geral</b>	<b>12</b>
<b>1.1.2</b>	<b>Objetivos Específicos</b>	<b>12</b>
1.2	DEFINIÇÃO DO ESCOPO	12
1.3	ORGANIZAÇÃO DO TRABALHO	13
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	<b>14</b>
<b>3</b>	<b>BACKGROUND</b>	<b>16</b>
3.1	CHATBOTS	16
3.2	TECNOLOGIAS PARA CHATBOTS	17
<b>3.2.1</b>	<b>Rasa</b>	<b>17</b>
<b>3.2.2</b>	<b>DialogFlow</b>	<b>17</b>
<b>3.2.3</b>	<b>IBM Watson</b>	<b>18</b>
3.3	UTILIZAÇÕES DE <i>CHATBOTS</i> NA ATUALIDADE	18
3.4	PROCESSAMENTO DE LINGUAGEM NATURAL	19
<b>3.4.1</b>	<b>spaCy</b>	<b>20</b>
<b>3.4.2</b>	<b>Word Embeddings</b>	<b>21</b>
<b>4</b>	<b>DESENVOLVENDO UM CHATBOT</b>	<b>23</b>
4.1	PIPELINE DO RASA	24
4.2	RASA CORE	24
4.3	UNIDADE DE PROCESSAMENTO DE LINGUAGEM NATURAL (NLU)	26
4.4	INTERFACE COM USUÁRIO	28
4.5	DOMÍNIO SOBRE VACINAS DO COVID-19	28
<b>5</b>	<b>DESENVOLVENDO UM CHATBOT PARA RESPONDER DÚVIDAS SOBRE AS VACINAS CONTRA COVID-19</b>	<b>31</b>
5.1	LEVANTAMENTO DE DÚVIDAS SOBRE AS VACINAS	31
5.2	TREINAMENTO DA NLU	32
5.3	ENGENHARIA DO CONHECIMENTO	33
5.4	DESENVOLVIMENTO DAS <i>STORIES</i>	39
<b>6</b>	<b>AVALIAÇÃO DO CHATBOT DESENVOLVIDO</b>	<b>48</b>
6.1	AVALIAÇÃO DA NLU	48
6.2	AVALIAÇÃO DAS <i>STORIES</i>	51
<b>7</b>	<b>CONCLUSÃO E TRABALHOS FUTUROS</b>	<b>53</b>
7.1	TRABALHOS FUTUROS	53
	<b>REFERÊNCIAS</b>	<b>54</b>

## 1 INTRODUÇÃO

Uma das principais características de uma sociedade globalizada é o fácil acesso a informação. Com o acesso cada vez mais rápido e fácil a internet, ter na palma da mão a informação pela qual se busca se tornou um corriqueiro hábito de parcela considerável da sociedade. Não importa o tema ou o assunto do debate, a quantidade de dados, informações e estatísticas sobre o mesmo será enorme. Em paralelo á facilidade pelo acesso a informações, existem eventos que trazem ainda mais necessidade pela procura por dados e informações. A pandemia da covid-19 é um desses eventos. Um dos motivos pelo qual a pandemia despertou uma necessidade enorme por estar bem informado foi a sucessão de eventos que ela desencadeou. Após os primeiros casos confirmados vieram os *lockdowns*, hospitais por todo o mundo lotados em virtude dos altos números de casos e óbitos registrados, eventos cancelados e adiados, sejam no Brasil ou no exterior, comércios fechados e o surgimento de variantes do vírus.

A pandemia monopolizou as atenções por meses. Virou parte do cotidiano, mas ao mesmo tempo que ela trouxe várias dúvidas, trouxe uma certeza: o cenário melhoraria com o surgimento das vacinas. Após laboratórios e farmacêuticas começarem a fabricar e distribuir as vacinas mundialmente, surgiram ainda mais informações, desta vez não sobre a pandemia, mas sim sobre a ferramenta mais essencial para o controle da mesma, as vacinas. As dúvidas que permeavam o coronavírus agora permeavam as vacinas. Seriam elas eficazes e seguras? Estas e várias outras perguntas podem ser feitas sobre as mesmas. Dúvidas e informações em demasia para uma escassez de certezas.

Visando explorar esse domínio de aplicação e o estudo de tecnologias *chatbots*, nesse trabalho desenvolvemos um *chatbot* que tem como objetivo sanar dúvidas e responder perguntas sobre as vacinas da COVID-19.

De acordo com (ADAMOPOULOU; MOUSSIADES, 2020), *chatbots* são ferramentas extremamente poderosas âmbito de comunicação entre o ser humano e o software. Poucas ferramentas são capazes de realizar uma comunicação e interação entre o ser humano e a máquina quanto o *chatbot*, isso se deve ao fato da ferramenta conseguir manter uma conversa com o usuário, emulando uma conversa entre duas pessoas. Não apenas a comunicação é satisfatória, como a interação entre *chatbot* e usuário também é de um alto nível de satisfação. Isso ocorre pois um *chatbot* é programado e desenvolvido para sanar dúvidas, responder mensagens e realizar tarefas simples, como pode ser observado em (ADAMOPOULOU; MOUSSIADES, 2020). Tal troca de mensagens tem como objetivo primário, satisfazer as intenções do usuário, independente de quais sejam e da maneira a qual ele se comunica. Seja usando uma linguagem formal ou não, seja com erros de digitação, gramática ou ortografia, um *chatbot* deve conseguir entender a mensagem e manter uma conversa com o usuário. Outra vantagem é a sua fácil escalabilidade, isto é, fazer a manutenção ou alteração de dados e componentes de um *chatbot* é simples e rápido,

muitas vezes sem depender de agentes externos, como o acesso a internet, por exemplo. Os *chatbots* então, podem ser considerados robustos, independentes, fáceis de atualizar e modificar e amplamente eficazes no desempenho de suas funções. Dito isso, fica mais fácil de compreender o por que da escolha de um *chatbot* em detrimento de uma *webpage*, por exemplo (SHARMA; JOSHI, 2020). O cenário da pandemia da covid-19 contempla várias dúvidas e incertezas sobre o futuro, mas com o auxílio de uma ferramenta como um *chatbot*, dúvidas podem ser sanadas e perguntas sobre o contexto das vacinas contra a covid-19 podem ser respondidas.

## 1.1 OBJETIVOS

Nas seções abaixo estão descritos o objetivo geral e os objetivos específicos deste trabalho de conclusão de curso.

### 1.1.1 Objetivo Geral

Realizar um estudo sobre tecnologia de *chatbots* e validar esse estudo com o desenvolvimento de um *chatbot* capaz de responder dúvidas sobre as vacinas da covid-19, de maneira eficiente e objetiva.

### 1.1.2 Objetivos Específicos

- Realizar um estudo sobre tecnologias utilizadas na implementação de *chatbots*;
- Realizar um estudo sobre o domínio de vacinas da covid-19, coletando informações de fontes confiáveis para construir a base de dados a ser utilizada pelo *chatbot* a ser desenvolvido;
- Desenvolver um *chatbot* capaz de responder questões relacionadas sobre as vacinas da covid-19;
- Avaliar empiricamente o *chatbot* desenvolvido.

## 1.2 DEFINIÇÃO DO ESCOPO

O desenvolvimento do *chatbot* ocorreu durante o período da pandemia da covid-19, entre setembro de 2021 e fevereiro de 2022. Após o término da implementação, surgiram novas variantes da covid-19, entre elas a delta e a omicron. Também começaram a surgir as doses de reforço das vacinas e novos estudos sobre as mesmas eram publicados semanalmente. No entanto, informações sobre as variantes novas e as doses de reforço não foram contempladas na implementação por falta de informações e tempo de desenvolvimento.

### 1.3 ORGANIZAÇÃO DO TRABALHO

Na seção 2, são descritos alguns trabalhos e artigos que tem como foco o uso de tecnologias de *chatbot* e as conclusões que puderam ser observadas de cada um desses trabalhos e uma justificativa para o desenvolvimento desse trabalho. Na seção 3, é descrito como um *chatbot* funciona e suas aplicações atualmente, bibliotecas e ferramentas utilizadas no desenvolvimento do mesmo. Na seção 4, são descritas todas as etapas do desenvolvimento de um *chatbot*, todos os seus componentes e o domínio sobre as vacinas da covid-19. Na seção 5, são descritas as dúvidas levantadas e como o *chatbot* responde todas elas, as fontes de informação que serviram de base para o desenvolvimento e como o *chatbot* interage com o usuário em cada cenário. De maneira empírica, são feitos testes que avaliaram os resultados do *chatbot*. Os resultados numéricos e gráficos, e como os testes são realizados estão descritos na seção 6. Por fim, os trabalhos futuros e a conclusão do trabalho desenvolvido estão presentes na seção 7.

## 2 TRABALHOS RELACIONADOS

Em (ROSSATO, Danúbia Miorando *et al.*, 2020a), os autores apresentam um estudo na Universidade de Passo Fundo como objeto de entender o engajamento de trabalhadores da universidade sobre o uso de chatbots para atendimentos. Esse estudo traz a seguinte pergunta “Qual o engajamento dos trabalhadores em um projeto voltado a implementação de um chatbot para atendimento a usuários?”. Os principais objetivos são: Descrever a implementação do chatbot, descrever os elementos e recursos na implementação do chatbot, identificar a percepção dos trabalhadores com relação aos resultados obtidos assim como o engajamento dos mesmos. Os dados utilizados para elaboração desse artigo foram coletados por meio de entrevistas e análises documentais. Se chegou a conclusão que o objetivo inicial, de constatar o engajamento dos trabalhadores na implementação do chatbot, foi alcançado, apesar das dificuldades encontradas, como falta de recursos e a baixa demanda de trabalho. Outra conclusão retirada do estudo é que, caso uma instituição de ensino deseje que os trabalhadores possuam engajamento máximo ao desempenhar uma atividade, ela deve fornecer as condições adequadas para tal, bem como uma divisão de tarefas justa e ideal para cada trabalhador, caso contrário, o engajamento pode ser afetado e conseqüentemente, o resultado final da atividade.

Em (SOUZA; RIBEIRO; BRETERNITZ, 2019), os autores descrevem os processos de desenvolvimento de chatbots, desde a definição de sua estrutura, onde seria aplicado, linguagens de programação e interface que serão implementadas, ao mesmo tempo que se desenvolve um chatbot que tem como finalidade fornecer informações e estatísticas sobre uma base de dados sobre filmes. Essas informações variam de nota do filme em sites de avaliação, ficha técnica, gênero, ano, bilheteria, melhores filmes de cada ano, filmografia de atores e diretores entre outros. Também foi detalhado o processo de desenvolvimento do chatbot integrado em um site.

Em (MOURA *et al.*, 2019), os autores fazem um mapeamento estatístico pertinente sobre as diversas aplicações e tecnologias aplicadas referente a chatbots. Esse mapeamento consiste em um levantamento em diversos artigos publicados no portal CAPES. Foram feitos a partir desses levantamentos e pesquisas, diversas constatações. Como por exemplo, países onde chatbots são utilizados com maior frequência, crescimento no número de patentes e publicações científicas sobre chatbots, que inclusive, conta com gráficos que apontam o número de patentes por países, assim como o número de patentes por ano. Uma das conclusões finais do artigo é que existe um aumento considerável a partir de 2016, no interesse das indústrias sobre essa tecnologia. Os pesquisadores também puderam concluir que houve um significativo aumento no número de patentes relacionadas a chatbots, em diversos países, e que tal aumento se deve a crescente difusão da tecnologia de chatbots em diversos países.

Em (KUYVEN *et al.*, 2018), os autores fizeram um levantamento de aplicação

de chatbots em diversas áreas do ensino, em bases nacionais e estrangeiras. Além de fazer uma revisão literária, esse artigo também responde questões como "objetivo de utilizar chatbots na educação?", "quais as principais técnicas e tecnologias utilizadas na implementação de chatbots educacionais?" entre outras. Cada pergunta é respondida de maneira satisfatória. Com relação aos levantamentos, se destacam as áreas de ensino que mais utilizam chatbots educacionais e em quais níveis educacionais chatbots estão mais presentes. O nível educacional que apresentou o maior índice de utilização de chatbots foi o ensino superior, principalmente na área da ciência da computação. Também foi possível apurar as maiores dificuldades em difundir o uso de chatbots; uma base de conhecimento extensa e dificuldade de implementar fluxos de conversa entre o software e o aluno.

Pelo nosso conhecimento, o chatbot desenvolvido é a primeira iniciativa de desenvolver um chatbot que responde dúvidas relacionadas as vacinas da COVID-19 fato que estimulou o desenvolvimento do chatbot, pois uma das motivações é prover uma fonte de informações confiáveis sobre as vacinas na qual o usuário possa interagir como quiser com a máquina. Embora nosso objetivo não seja torná-lo público inicialmente.

## 3 BACKGROUND

### 3.1 CHATBOTS

*Chatbot* pode ser definido como um software que tem como finalidade manter uma conversa com um usuário humano. Essa conversa pode ser via aplicativo de mensagens, sites, call centers entre outros. As aplicações de um chatbot são variadas. Podem ter como funcionalidades sanar o maior número possível de dúvidas sobre um assunto em específico, auxiliar o usuário com dificuldades, ou ainda, auxiliar o usuário em pedir uma refeição via aplicativo de mensagens, entre outras (PEREIRA; PINHEIRO, 2018).

Um dos maiores usos de tecnologias *chatbots* está presente em empresas e prestadoras de serviços que utilizam *chatbots* para manter contato com o usuário e o cliente, e essa crescente utilização da tecnologia se deve a alguns fatores (LUGLI; LUCCA FILHO, 2020):

- capacidade de oferecer resposta e suporte imediato ao usuário;
- redução na equipe de atendimento;
- disponibilidade 24 horas por dia, 7 dias na semana;

Um *chatbot* pode ser construído a partir de aprendizado de máquina (*machine learning*) ou estabelecimento de diretrizes. O *chatbot* que funciona mediante estabelecimento de diretrizes (ou regras) é capaz de responder perguntas e requisições que estejam dentro do conjunto de regras o qual ele foi inicialmente programado e treinado. Possui vocabulário limitado e definido por quem o desenvolveu. Por sua vez, o *chatbot* que utiliza *machine learning* consegue interpretar e aprender com as mensagens e requisições do usuário. Fica mais inteligente a medida que interage com o usuário e responde as perguntas e mensagens com mais precisão e é mais complexo de ser implementado (MOURA *et al.*, 2019).

Em geral, *chatbots* são utilizados nos seguintes setores (ROSSATO, Danúbia Miorando *et al.*, 2020b): atendimento, vendas e marketing. Na grande maioria dos casos, se utilizam *chatbots* com o intuito de se realizar uma compra ou aquisição de um serviço, mas os *chatbots* também são muito usados para tirar dúvidas do cliente sobre um determinado produto, por exemplo. *Chatbots* também são implementados para sugerir ao usuário um determinado item ou serviço. Nesse tipo de aplicação, com base no que é respondido, o *chatbot* realiza uma sugestão de compra para quem está conversando com ele. *Chatbots* são práticos e acessíveis. Na atualidade, não são apenas as grandes empresas e multinacionais que dispõem dessas tecnologias. Isso se deve aos frameworks *opensource* para o desenvolvimento de chatbots. Um dos mais famosos e utilizados é o Rasa, o qual será descrito em mais detalhes na Seção 3.2.1.

Outro motivo pelo qual empresas tem investido cada vez mais em *chatbots* é a aproximação com o cliente. As empresas acreditam que *chatbots* as tornam mais próximas

do cliente, além de proporcionar a vantagem de ser mais eficiente e rápido o atendimento se comparado ao atendimento por telefone, via *call center*. Por fim, outra vantagem dos *chatbots* é a não necessidade de se realizar o download de cada aplicativo de cada loja ou serviço. Economizam memória no aparelho de cada usuário e são mais práticos (ROSSATO, Danúbia Miorando *et al.*, 2020b).

## 3.2 TECNOLOGIAS PARA CHATBOTS

### 3.2.1 Rasa

Como pode ser observado em (SHARMA; JOSHI, 2020), uma das ferramentas mais usadas para o desenvolvimento de *chatbots* é o Rasa<sup>1</sup> *Open Source*. O qual possibilita a implementação de *chatbots* nos sistemas operacionais *Windows*, *Linux* e *macOS*. Pode ser importado para sites e aplicativos de mensagens, como *Telegram* e *WhatsApp*. Utilizando bibliotecas como *spaCy* e *Tensorflow*, o Rasa possibilita o desenvolvimento de chatbots sofisticados, utilizando técnicas transparentes e flexíveis. Entre as principais características do Rasa estão:

- A capacidade de extrair significado de quaisquer mensagens em diversos idiomas.
- Possui suporte para estruturas que possuem aprendizado de máquina e os que operam utilizando diretrizes.
- É altamente customizável e possui ampla gama de idiomas em seus domínios.
- É capaz de utilizar as regras de negócio de uma empresa, bem como suas diretrizes em mensagens e conversas fluídas e agradável aos usuários.
- Possui suporte a diversos aplicativos de mensagens, como *WhatsApp*, *Telegram*, *Facebook* e *Google Home*.
- Por fim, possui uma documentação muito ampla (embora muitas vezes confusa devido as alterações de versionamentos) que facilita o desenvolvimento dos chatbots.

### 3.2.2 DialogFlow

O DialogFlow é um serviço provido pelo *Google*. É definido como uma inteligência artificial de conversação com agentes virtuais. Tal qual o Rasa, opera tanto com inteligências artificiais como com bancos de dados previamente estabelecidos. Uma de suas principais diferenças em relação ao Rasa, é o uso de um fluxo visual interativo que tem como o enfoque o uso de vários desenvolvedores simultaneamente. Uma das principais empresas que utilizam o serviço do *DialogFlow* é a companhia aérea *Malaysian Airlines*,

---

<sup>1</sup> <https://rasa.com/>

que usa um *chatbot* como ferramenta para pesquisa de preços e compra de passagens. No entanto, o uso de todos os recursos disponíveis pelo *DialogFlow* é limitado a um número de vezes que se utiliza o programa. Essa limitação ocorre em sua versão gratuita, o que torna o uso dessa ferramenta pouco produtivo, uma vez que existem opções gratuitas mais vantajosas. Por se tratar de um trabalho acadêmico, foi optado em utilizar o Rasa em função de possuir um custo uma documentação mais satisfatória e oferecer todas as ferramentas necessárias para o desenvolvimento do chatbot (INÁCIO; CANATO, 2020).

### 3.2.3 IBM Watson

Outra ferramenta comum no desenvolvimento de chatbots é o IBM Watson. Usado por diversas empresas no atendimento automático aos seus clientes. Entre as companhias que usam esta ferramenta, destaca-se a companhia de telecomunicações Vodafone. Possui as mesmas vantagens e funções do Rasa e do DialogFlow, no entanto, esbarra no mesmo problema do serviço do Google, a limitação de funções em suas versões gratuitas. Ainda que suas versões pagas possuam vantagens como uma interface ao usuário limpa e agradável, não foi considerado uma opção por ser restrito na disponibilidade de funções em sua versão gratuita.

## 3.3 UTILIZAÇÕES DE CHATBOTS NA ATUALIDADE

De acordo com (KANNAN; BERNOFF, 2019), o crescimento astronômico do uso de *chatbots* por parte das empresas tem algumas explicações. Notou-se que o uso de chatbots fideliza os clientes pois a interação com um chatbot não depende de um horário de atendimento de um *call center*, por exemplo. Isso fez com que as empresas ficassem mais eficazes e operantes. A integração com aplicativos de mensagens, como WhatsApp e Telegram<sup>2</sup> ajuda a popularizar a ferramenta, principalmente no Brasil, um dos países onde aplicativos de mensagem como WhatsApp possui mais usuários. Juntando o alto número de usuários de aplicativos de mensagens com a portabilidade de chatbots nesses aplicativos, o número de usuários de chatbots sobe por consequência. Para se ter ideia do crescimento dessa ferramenta entre os brasileiros, o Brasil possui 3 por cento de todos sites que utilizam algum tipo de chatbot no mundo (SIENKIEWICZ, 2022). Em decorrência da portabilidade entre aplicativos de mensagens e chatbots, não apenas grandes multinacionais tem usufruído desta ferramenta. Micro e pequenas empresas também tiram proveito dessa interação entre aplicativos e chatbots para atrair clientes e consumidores. Isso ocorre pois essa portabilidade é de fácil implementação e manutenção, além de possuir um baixo custo.

<sup>2</sup> <https://www.proxima.com.br/home/proxima/how-to/2020/01/10/uma-breve-historia-do-mercado-de-chatbots-no-brasil.html>

Os assistentes virtuais tem ganhado espaço dentro das empresas, não apenas quando o intuito é conversar com o usuário. É cada vez mais comum os setores de recursos humanos utilizarem os assistentes virtuais para realizar tarefas como agendamento de férias e emissão de documentos. Esses assistentes virtuais também servem como canais de denúncia em casos de assédio dentro da empresa e de guia para auxiliar novos empregados com as políticas da empresa. O ramo da medicina também tem obtido vantagem a partir de chatbots. Hospitais e consultórios usam essa ferramenta para agendar consultas e realizar de maneira mais eficaz, a coleta de dados para pré-exames. Um exemplo é o WHO Health Alert da Organização Mundial da Saúde que envia informações sobre a disseminação do Covid-19 para celulares e smartphones via aplicativos de mensagem, como o WhatsApp (ALERT, 2022).

Um bom exemplo de implementação de *chatbot* é visto na empresa Casas Bahia. A rede de varejos utiliza *chatbots* em sua página no *Facebook* em datas de grande volume de vendas, como dia do consumidor, *Black Friday* e Natal. O *chatbot* que ficou conhecido entre os clientes como “Baianinho”, e oferecia diversas ofertas de produtos ao usuário, de diversas categorias. Durante a *Black Friday* de 2018, 51.4 mil pessoas interagiram com o *chatbot*, totalizando quase 2 milhões de mensagens (TUCUNDUVA, 2022).

O aplicativo *Uber* é outro caso de sucesso com relação ao uso de *chatbots*. Uma corrida pode ser marcada via *Facebook*, em contato com o bot na página oficial do *Uber* na rede social (TUCUNDUVA, 2022).

Estratégia similar utiliza a rede americana *Pizza Hut*. Via chat no *Facebook Messenger*, o usuário pode realizar um pedido conversando com o *chatbot* da empresa. O cliente também pode vincular sua conta na rede social com o *Pizza Hut*, que pode consultar o histórico de pedidos de cada cliente e oferecer ofertas personalizadas (PIZZAHUT, 2017).

Independente da ferramenta, da técnica ou do propósito utilizado para o desenvolvimento de um *chatbot*, existem denominadores em comum entre eles. Um deles, é o Processamento de Linguagem Natural. Sem essa técnica, seria impossível desenvolver um *chatbot* de maneira adequada. Tudo o que essa técnica aborda, método de funcionamento entre outras características, serão abordadas em detalhes na seção a seguir.

### 3.4 PROCESSAMENTO DE LINGUAGEM NATURAL

Processamento de Linguagem Natural, ou PLN, é uma sub-área da Inteligência artificial cujo objetivo e função é compreender e assimilar linguagem natural nos quesitos composição e compreensão da linguagem humana, seja por escrita ou por fala. Pode ser comparado a uma criança que está aprendendo suas primeiras palavras. Quanto mais informações e estímulos receber, mais facilmente assimila a linguagem com a qual está tendo contato (FINATTO; LOPES; SILVA, 2015). No caso das máquinas, o aprendizado depende do modo como foi programada e a frequência com a qual é treinada. PLN pode ser dividido em 7 níveis, cada um representado uma área da linguística. São eles:

1. Semântico: Compreensão de uma frase inteira;
2. Sintático: Composição de uma frase;
3. Fonológico: Interpretação dos sons das palavras;
4. Morfológico: Natureza e composição das palavras;
5. Discurso: Estudo de um texto inteiro;
6. Lexical: Interpretação de cada palavra;
7. Pragmático: Interpretação do que pode estar nas entrelinhas de uma frase.

Em geral, PLN apresenta abordagens com regras linguísticas precisas (abordagem simbólica), modelos matemáticos (abordagem estatística), que conecta modelos estatísticos com outras formas de representar o conhecimento (abordagem conexionista) e abordagens híbridas que são as teorias citadas anteriormente conectadas entre si. Interpretação de frases ambíguas e detecção de palavras fora de ordem em uma frase são algumas dentre as várias aplicações do PLN (ALLEN, 2003).

As vantagens e facilidades contidos no PLN logo passaram a ser implementadas em outras tecnologias, incluindo chatbots. Isso ocorre pela capacidade do PLN auxiliar softwares a compreenderem e interpretarem a linguagem humana. Uma vez que chatbots estão sempre interagindo e mantendo diálogos com usuários, a implementação de PLN em chatbots foi fundamental para que não apenas chatbots, mas ferramentas como assistentes pessoais, como a Siri e a Alexa, também tirassem proveito das funções do PLN. O PLN é essencial na interpretação de mensagens com erros gramaticais, ou de ortografia, de maneira que um chatbot possa entender o que esteja sendo dito pelo usuário, ainda que de maneira equivocada ou sem coerência. Em casos mais complexos e elaborados, chatbots dotados de PLN são até capazes de reconhecer o estado emocional do usuário, tamanha a compreensão que o PLN é capaz fazer da linguagem usada pelo ser humano (PEREIRA; PINHEIRO, 2018).

Entre as bibliotecas utilizadas para processamento de linguagem natural, a mais recomendada e utilizada para implementações de chatbots, inclusive recomendada pelo próprio site do Rasa, é o spaCy. Introduziremos o biblioteca spaCy na próxima seção.

### 3.4.1 spaCy

O spaCy apresenta uma série de modelos pré-treinados de processamento de linguagem que são capazes de realizar diversas funções como tokenização, reconhecimento de entidades, transformar palavras em vetores, entre outras<sup>3</sup>.

---

<sup>3</sup> Mais informações em <https://spacy.io>

Como citado anteriormente, no processo de execução de um chatbot, existe o processo de tokenização, onde um texto é fragmentado em frases e as frases fragmentadas em palavras. No entanto, existem palavras que são consideradas ‘*stop words*’. Alguns exemplos de stop words são, “o”, “era”, “isso”, assim como sinais de pontuação, como ponto final e vírgula. Essas stop words são análogas a um ruído. O processo de remoção de ruído de um texto ou uma frase, é chamado de limpeza de texto e pré-processamento. O pré-processamento é fundamental para a performance de PLN. Stop words não adicionam significado a um texto e podem prejudicar o resultado de PLN. Para que isso não aconteça, é necessário desconsiderar as stop words de um texto. Para reconhecer as stop words, o spaCy possui atributos como `token.is_stop`, `token.is_punct` e `token.is_space`<sup>4</sup>. Outra vantagem do pré-processamento, é a diminuição do tamanho do arquivo e conseqüentemente, redução do custo computacional para executá-lo.

Outra ferramenta inclusa no spaCy é a lematização. Essa ferramenta é responsável por um processo no qual palavras derivadas são traduzidas para sua forma base. Por exemplo, as palavras “corri”, “correu” e “correndo” são traduzidas para sua forma primária, nesse caso, a palavra “correr”. A vantagem da lematização, é a redução no número de tokens dentro do processamento da PLN.

O spaCy também possui uma ferramenta cuja função é atribuir cada palavra dentro de um texto, um identificador único e exclusivo. Isso funciona muito bem em casos onde palavras se repetem diversas vezes. Cada palavra recebe um identificador único e exclusivo, localizado no arquivo `StringStore`. O `stringStore` é um dicionário que mapeia cada palavra e atribui a ela um valor. Essa ferramenta ajuda a poupar memória, uma vez que cada palavra possui seu próprio identificador, não havendo a necessidade de atribuir um identificador diferente para cada vez que uma mesma palavra se repetir em uma sentença ou texto. Outro método importante é o `pos_attribute`. Esse atributo é o encarregado em reconhecer a classe gramatical de cada palavra contida em uma sentença.

Por fim, o componente mais importante do spaCy, o vetor de palavras. Vetores de palavras são representações numéricas das palavras de um texto ou de uma sentença. Através da função `token.vector_norm` que atribui cada valor para os vetores de palavras. Para verificar o quanto duas sentenças são similares, é utilizado uma métrica conhecida como “*cosine similarity*”. Essa fórmula considera a distância angular entre dois vetores num espaço tridimensional. O resultado varia entre 0 e 1, quanto mais próximo de 1, maior a similaridade entre duas sentenças. Tais vetores multidimensionais são conhecidos como word embeddings, que serão introduzidos na próxima seção.

### 3.4.2 Word Embeddings

Word Embeddings pode ser definido como um conjunto de técnicas onde palavras são representadas matematicamente como um vetor (NAILI; CHAIBI; GHEZALA, 2017).

<sup>4</sup> <https://spacy.io/usage/spacy-101>

Esses vetores possuem tamanho fixo e armazenam o contexto e o significado de um texto ou enunciado. A representação de cada palavra em um campo multidimensional é chamada de *embedding space*. Na abordagem do *word embeddings*, palavras são representadas em escalas que representam informações sobre seu significado. Por exemplo, a palavra “enfermeira” pode ser representada numa escala de gênero que vai de -1 a 1, quanto mais perto de -1, mais feminina a palavra é e quanto mais perto de 1, mais masculina. No entanto, apenas uma classificação de gênero de uma palavra não é o suficiente para representá-la com clareza. Neste caso, são atribuídas mais valores ao vetor da palavra, cada um representando o valor em uma escala distinta. Para que todas as palavras contidas em um texto tenham seus vetores preenchidos, são utilizadas técnicas de *machine learning*, como o Word2Vec e a similaridade entre vetores.

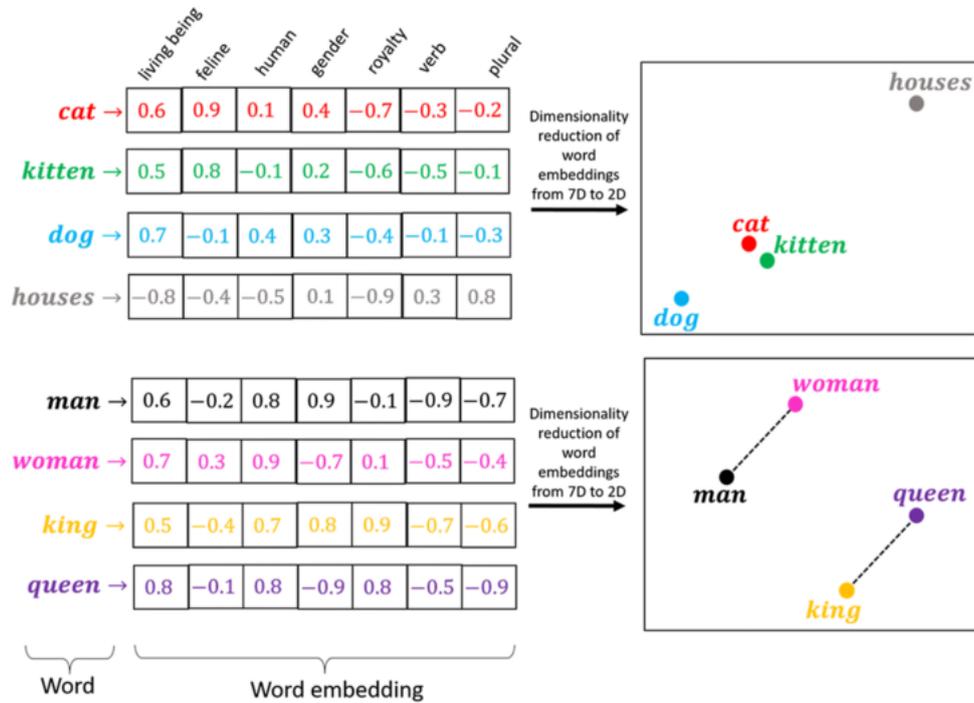


Figura 1 – Word2Vec

## 4 DESENVOLVENDO UM CHATBOT

Rasa é um *framework open source* cuja função primária é a implementação de *chatbots*. No ramo das ferramentas utilizadas para desenvolver *chatbots*, o rasa não é o único, mas é um dos *frameworks* mais utilizados. O rasa apresenta vantagens e diferenciais em relação as demais ferramentas de desenvolvimento de *chatbots*. A primeira é de oferecer uma vasta gama de utilidades sem custo. Diferente do DialogFlow, por exemplo, que apresenta recursos limitados gratuitamente. A segunda vantagem do Rasa, é de ser totalmente customizável. Possui componentes como o spaCy, descrito na seção anterior, que traz suporte a diversos idiomas, incluindo o português. O spaCy também é responsável por outros recursos profícuos, como o modelo de linguagem criado através de um corpus de notícias em português, o "pt\_core\_news\_sm", que foi de grande utilidade na implementação do *chatbot*. Outra vantagem do Rasa, é o suporte a aplicativos de mensagens como Telegram e Whatsapp, e HTTP API<sup>1</sup>.

Também, todas as etapas de construção do *chatbot* utilizando o Rasa podem ser realizadas offline, sem a necessidade de atualizações do projeto para a nuvem. Essas, entre outras, características foram cruciais na escolha do *framework*.

Para um melhor entendimento do Rasa, é necessário compreender algumas palavras-chaves que utilizaremos nessa seção. São elas:

- **Intenções (Intents):** São o resultado de um processo de classificação de frases em linguagem natural. Uma intenção descreve significado de uma mensagem. São utilizadas nas tomadas das decisões de um *chatbot*. Dependendo da mensagem enviada pelo usuário e na intenção extraída da mesma, o *chatbot* realizará sua tomada de decisão.
- **Entidades (Entitys):** São trechos de frases que ajudam o *chatbot* a compreender especificamente do que o usuário está falando e tomar decisões. Por exemplo, caso o usuário digite uma mensagem dizendo que gostaria de comprar um celular de uma marca específica, a entidade reconhecida seria o nome da marca desejada.
- **Ações (Actions):** São ações que o *chatbot* pode executar e descrevem como responder uma iteração do usuário, podem abranger consultas externas, como site, banco de dados, comunicar com dispositivos externos, etc.
- **Respostas (Responses):** Respostas que o *chatbot* pode comunicar ao usuário, normalmente no formato textual.

---

<sup>1</sup> <https://rasa.com/docs/rasa/http-api/>

## 4.1 PIPELINE DO RASA

O Rasa pode ser dividido em dois componentes principais, Rasa NLU (*Natural Language Understanding*) e Rasa Core. O NLU é encarregado de extrair entidades e intenções das mensagens enviadas pelo usuário. Para realizar tal tarefa, o NLU utiliza técnicas de *machine learning* para analisar o conteúdo de uma mensagem. O Rasa core, por sua vez, é responsável por gerenciar a conversa entre o *chatbot* e o usuário. É o componente responsável por fornecer caminhos de tomadas de decisão em que as ações e mensagens que devem ser expressas pelo *chatbot*, a medida que o diálogo flui. Esses dois componentes serão detalhados nas seções a seguir:

## 4.2 RASA CORE

Como citado na seção 3.2.1, existe um arquivo no rasa chamado `config.yml`. É nesse arquivo onde são configurados os componentes do pipeline do rasa. Existem diversos componentes que podem ser aplicados para o desenvolvimento de um *chatbot*. Todos os componentes apresentam funções distintas e que por muitas vezes se completam. No entanto, não são todos os componentes que apresentam utilidade no desenvolvimento do *chatbot* descrito neste trabalho. Nesta seção, citaremos os componentes do pipeline do rasa que foram utilizados e a função de cada um deles.

No arquivo `domain.yml`, é especificado o domínio em que o *chatbot* será desenvolvido, incluindo as intenções do usuário a serem reconhecidas (*intents*), ações a serem executadas (*actions*), os slots de memória que serão utilizados, as mensagens que serão mostradas ao usuário e toda e qualquer ferramenta que precise ser implementada, como por exemplo os forms. Slots de memória são componentes não obrigatórios no desenvolvimento de um *chatbot*. São informações enviadas pelo usuário que são armazenadas pelo *chatbot* e que baseado do conteúdo da informação, o *chatbot* toma ações distintas no decorrer da conversa com o usuário. Outra ferramenta disponível no Rasa são as rules/regras. Tal ferramenta determina partes de conversas que devem estar presentes em todas as mensagens, independente das *stories* nas quais ela está contida. As ações tomadas pelo *chatbot* são diversas. Podem ser simples como fornecer uma resposta simples ao usuário, ou algo mais elaborado, como dar sugestões de compra de algum produto, checar a previsão do tempo ou pedir as notícias do dia. Esse tipo de ação mais elaborada é conhecida como *custom action*, e deve ser devidamente implementada no arquivo `actions.py`. No desenvolvimento de um *chatbot* que apenas fornece respostas e faz questionamentos, não são necessárias *custom actions*.

O primeiro componente do pipeline, é o mais famoso e crucial entre todos os componentes de um pipeline: O *tokenizer*. Sua função é fragmentar uma frase em pedaços menores de texto. Esses pedaços menores são conhecidos como *tokens*. Esses *tokens* depois serão processados por outras partes do pipeline com a finalidade de extrair entidades e

ações, por exemplo. Outra parte do pipeline é o `CountVectorsFeaturizer`. Esse componente tem como função gerar um *bag of words* para palavras, intenções e ações. *Bag of words* é uma ferramenta muito utilizada em pré processamento de dados, na qual, cada palavra contida em uma frase, ou em um texto, é contada e o número de vezes em que tal palavra se repete, fica armazenado em um vetor. Quanto maior a frequência de uma palavra em um texto, maior a sua relevância e, conseqüentemente, sua importância para o treinamento e execução de um modelo.

O DIET (*Dual Intent and Entity Transformer*) é outro componente do pipeline do `rasa`. Utilizado para classificação de intenção e extração de entidades. Baseado na iteração com o usuário, o DIET classifica a qual das intenções presentes no modelo a mensagem se classifica. Essa predição é necessária para extrair entidades da mensagem do usuário e, conseqüentemente, responder a iteração do usuário da maneira mais satisfatória possível. O algoritmo *whiteboard* é utilizado pelo DIET para realizar essas funções. Em seguida, o componente do pipeline chamado `EntitySynonymMapper` mapeia e reconhece sinônimos na extração de uma entidade.

Uma conversa entre o usuário e um *chatbot* possui como premissa que o *chatbot* responda as mensagens e perguntas fornecidas pelo usuário da maneira mais correta, precisa e coerente possível. Como podemos observar, existem diversos componentes dentro de um *chatbot* que são responsáveis pelo processo de receber as mensagens do usuário e tratá-las de maneira que o *chatbot* saiba qual a intenção do usuário. Após reconhecer a intenção do usuário, é essencial que o *chatbot* saiba o que deve ser feito, qual ação executar e qual mensagem entregar como resposta. Nesse contexto, são definidas as *stories*. *Stories* funcionam como um roteiro, um guia, no qual o *chatbot* se baseia para saber quais ações deve realizar ou quais mensagens deve emitir de acordo com as intenções e mensagens proferidas pelo usuário. De acordo com cada intenção reconhecida pelo *chatbot*, o mesmo deve desempenhar uma ação de acordo com as *stories*. As *stories* são essenciais para que a conversa siga um fluxo coeso e coerente e para que o usuário tenha suas mensagens devidamente respondidas. O tamanho de cada *storie* e a quantidade das mesmas fica a critério do desenvolvedor, variando com a complexidade do *chatbot*.

Por fim, existem as *policies*. As *policies* são definidas e configuradas no arquivo `config.yml` e tem como utilidade definir os rumos da conversa entre o usuário e o *chatbot* e quais ações serão tomadas pelo *chatbot* após o reconhecimento da intenção do usuário e a extração das entidades contidas naquela interação. Para o desenvolvimento deste artigo, foi utilizada a `MemoizationPolicy`, a `TEDPolicy`, a `MappingPolicy` e a `FormPolicy`. Essa *policy* é de suma importância em *chatbots* que possuem várias *stories*, pois é ela que verifica se as conversas entre o usuário e o *chatbot* estão seguindo o caminho descrito pelas *stories*. Caso a troca de mensagens seja a mesma descrita em uma das *stories* definidas, a `MemoizationPolicy` irá fazer com que o *chatbot* tome as ações pré-estabelecidas nas *stories*. A `MemoizationPolicy` é responsável pelo andamento correto das *stories* e conseqüentemente,

```
intent: greet
examples: |
- oi
- ola
- bom dia
- boa tarde
- boa noite
- como vai?
- oie
- saudações
- oi, bom dia
- oi boa tarde
- ola boa noite
- tudo bem?
```

Figura 2 – Exemplos de entrada na NLU.

do *chatbot*.

### 4.3 UNIDADE DE PROCESSAMENTO DE LINGUAGEM NATURAL (NLU)

Para o desenvolvimento de um *chatbot* são utilizadas técnicas de Processamento de Linguagem Natural (PLN). PLN pode ser definido como área de conhecimento que visa a tradução da linguagem humana para o computador. PLN permite que a máquina entenda e interprete as mensagens providas pelo usuário. Essa comunicação pode ser feita tanto por meio de textos, quanto por meio de áudios. E não apenas entender o que está sendo dito pelo usuário, a máquina também fica apta a estabelecer um diálogo e responder as perguntas feitas pelo usuário. A assistente dos sistemas IOS, Siri, é um dos exemplos de aplicações de PLN. Como exemplo, suponha-se que esteja sendo desenvolvido um *chatbot* com diversas intenções, e uma delas é a intenção de saudação. Existem diversas maneiras de se saudar uma pessoa, como podem se observar na Figura 2.

Se pode constatar que existem diversas maneiras de se dizer uma mesma coisa, no caso, diversas maneiras de expressar uma saudação. O Rasa NLU então, é responsável por reconhecer saudações distintas como uma única intenção, no caso, a intenção de saudação. Para que isto ocorra, o Rasa NLU treina um classificador que extraia de sentenças e frases em linguagem natural as intenções e realize a extração de entidades. Como consequência, se um usuário digitar "olá meu amigo", o *chatbot* será capaz de retornar a intenção saudação. A frase "olá amigo" não está presente nos exemplos de treinamento contidos na Figura 2. Isso ocorre em decorrência da generalização do algoritmo de *machine learning*, que reconhece a similaridade de mensagens com intuito de fazer a classificação e extração de intenções. Tal processo ocorre mesmo com frases não pertencentes aos exemplos de treinamento. No entanto, para que isso ocorra, são necessários vários exemplos de treinamento, para que a extração de intenções seja a mais correta e precisa possível.

Segundo (KANNAN; GURUSAMY *et al.*, 2014), são duas etapas que compõe um processo de processamento de linguagem natural: o pré-processamento e o processamento.

O pré-processamento é composto pelas etapas de: Tokenização, remoção de stopwords (palavras como "a", "o", e "e"), remoção de numerais e símbolos, correção de erros ortográficos e de digitação e stemização (etapa onde uma palavra é reduzida ao seu radical) e lematização (etapa onde uma palavra é reduzida ao seu lema). Por fim, o processamento. Nesta etapa, existem 7 níveis. São eles: (i) Fonologia: etapa de reconhecimento de sons que formam palavras; (ii) Morfologia: divisão das palavras em morfemas e composição das palavras; (iii) Sintático: formação da sentença; (iv) Léxico: captação de significado a cada palavra individualmente; (v) Análise Semântica: extração do significado da sentença; (vi) Discurso: análise do significado na íntegra de um texto; e (vii) Pragmático: analisa se o resultado da análise semântica está correto.

Como citado anteriormente, um dos componentes cruciais para a elaboração de um *chatbot* é o pipeline. É neste componente que são definidas etapas que o *chatbot* irá utilizar para seu funcionamento. Cabe aos desenvolvedores decidir quais componentes e quais ferramentas irão utilizar para a implementação do *chatbot*. O Rasa no entanto, sugere um pipeline (em seus site, blogs e fóruns oficiais) que inclui modelos de linguagem disponíveis através do spaCy. As razões e vantagens do spaCy são diversas. A primeira, é de possuir pacotes em diversos idiomas, incluindo o português. Apesar de não possuir uma versão exclusiva para português brasileiro, isso não atrapalha o desenvolvimento do *chatbot*. Devido a imensa semelhança entre o português brasileiro e o europeu fornecido pelo spaCy, este pacote é satisfatório para o desenvolvimento do *chatbot*. Outra vantagem é a possibilidade de obter um pacote de tamanhos distintos. O vocabulário fornecido pelo spaCy pode ser obtido nos tamanhos pequeno, médio e grande. Fica a cargo dos programadores do *chatbot* escolher qual tamanho do vocabulário escolher para suprir suas necessidades. Outra vantagem são os componentes que interpretam frases parecidas e agrupam as mesmas em classes, ou seja, para frases semelhantes, a interpretação do *chatbot* será a mesma. Os componentes responsáveis por isso são o *SpacyFeaturizer* e o *MitieFeaturizer*.

Outro componente que contribui muito para o melhor funcionamento de um chatbot é o DIET. Tal componente fica encarregado de extrair intenções de frases e de selecionar que respostas o *chatbot* deverá fornecer ao usuário a partir das mensagens fornecidas pelo mesmo. O spaCy também oferece opções que utilizam e limitam o uso da placa gráfica do computador caso o usuário não deseje que a mesma seja utilizada integralmente no processo de treinamento do *chatbot*. Essas são algumas das ferramentas fornecidas pelo spaCy que foram contempladas no desenvolvimento do *chatbot*. No entanto, ele apresenta várias outras. Uma delas é *Profanity Filter*, que filtra mensagens de palavras de baixo calão. Com essa ferramenta, ofensas e xingamentos eventualmente fornecidas pelo usuário são censuradas pelo *chatbot* que irá ignorar as mesmas.

Entre os arquivos criados automaticamente pelo Rasa para o desenvolvimento do *chatbot* está o `config.yml`. Nesse arquivo estão presentes as políticas e os componentes

do pipeline que o desenvolvedor utiliza para construir o *chatbot*. Nesse arquivo também está presente o pipeline do Rasa, que foi detalhado na seção 4.1.

Um dos arquivos mais essenciais, é o `nlu.yml`. NLU é uma sigla em inglês para *Natural Language Understanding*. NLU pode ser traduzido como transformar mensagens do usuário em dados estruturados. Nesse arquivo `nlu.yml` ficam as intenções, onde seus nomes são definidos. Elas representam um conjunto de mensagens com o mesmo significado. A função do Rasa é prever a intenção certa quando o usuário enviar mensagens novas.

#### 4.4 INTERFACE COM USUÁRIO

O Rasa fornece uma interface de desenvolvimento, o Rasa X, que pode ser visualizado na Figura 3. Tal ferramenta acrescenta funcionalidades de desenvolvimento interativo, que proporciona um ambiente amigável para pessoas que não possuem experiência em desenvolvimento de software. Essa interface não foi utilizada no desenvolvimento do presente trabalho, considerando a experiência dos desenvolvedores na área de engenharia da computação.

Em termos de interface com o usuário, o Rasa possui portabilidade com diversos aplicativos de mensagem, como Telegram, Whatsapp e Facebook Messenger, além de possuir uma API Http. Tais interfaces são opcionais e não consideradas neste *chatbot*, uma vez que não faziam parte do escopo inicial do trabalho. Consideraremos interfaces com o usuário em nossos trabalhos futuros.

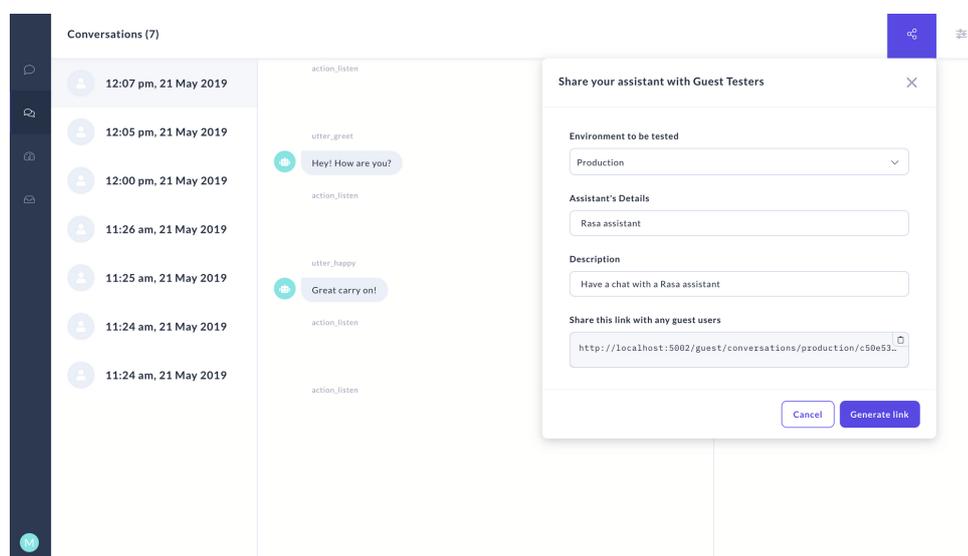


Figura 3 – Exemplo de interface de desenvolvimento do Rasa

#### 4.5 DOMÍNIO SOBRE VACINAS DO COVID-19

A pandemia da COVID-19 entrou para a história da humanidade por diversos fatores. Causou milhões de mortes, foi responsável direta e indiretamente por um prejuízo

financeiro global e mudou a rotina de todos. A pandemia apresentou consequências tão descomunais que muitas vezes, se tem apenas uma noção de seus desdobramentos e não valores, dados e números concretos.

Segundo a Organização Mundial da Saúde (ORGANIZATION, 2022), a pandemia causou aproximadamente 5,4 milhões de mortes com mais de 290 milhões de casos confirmados em todos os países do mundo, até o dia 4 de janeiro de 2022. Todas as vidas perdidas representam uma perda inestimável e irreparável para as famílias e amigos. A pandemia também causou danos financeiros. Os prejuízos financeiros causados pela pandemia, como estabelecimentos fechados ou com capacidade reduzida, eventos locais e globais cancelados ou adiados, como as Olimpíadas, etapas do mundial de Fórmula 1 e a Eurocopa, são muitas vezes difíceis de se estimar.

Em decorrência da pandemia, os prejuízos financeiros tem sido sentidos em diversas nações (BARRETT *et al.*, 2021). Nem mesmo o número exato de vítimas é possível afirmar. Não são todos os países que apresentam dados confiáveis para a OMS na hora de informar quantas vidas foram perdidas. Um cenário muitas vezes traumático e que traz muitas dúvidas, muitas vezes, mais dúvidas do que respostas. No entanto, se tem uma certeza: para voltar a normalidade, é necessário investir, acreditar e tomar as vacinas produzidas pelos laboratórios contra o coronavírus. Mas não é exatamente uma certeza unânime que as vacinas representam a luz no fim do túnel. No mundo inteiro, existem diversos movimentos e grupos de pessoas que propagam por diversos motivos (muitas vezes políticos), notícias falsas sobre as vacinas, as conhecidas *fake news*. Espalhar em redes sociais notícias falsas sobre as vacinas e sobre tratamentos ineficazes traz um outro tipo de prejuízo, a desinformação. Não são todas as pessoas que tem o bom hábito de verificar a veracidade de uma informação. Seja por preguiça ou por simplesmente acreditar na origem da mesma. Esse conflito entre informações muitas vezes causam dúvidas que não deveriam existir. Mas são fáceis de sanar.

É com esse objetivo que se chegou a ideia de desenvolver uma fonte de informações confiáveis e de fácil acesso, por meio de tecnologias *chatbot*. Todas as informações utilizadas para o desenvolvimento do *chatbot* são fornecidas pelos fabricantes das vacinas e/ou sites de periódicos e estudos altamente renomados e conceituados na comunidade científica. Tais dados estão sujeitos a alterações a medida que surgem novos imunizantes, estudos e doses extras das vacinas. No entanto, tais alterações são facilmente implementadas a medida que a estrutura de dados contida no *chatbot* é escalonável e de fácil atualização. A disseminação de *fake news* é um dos maiores problemas da sociedade contemporânea, traz prejuízos em diversos setores e segmentos. Numa era onde o acesso a informação é fácil e prático, o objetivo do *chatbot* é colaborar com essa facilidade a informação. Em prol da ciência e da informação, com bases de dados seguras e respaldadas pela comunidade científica, as informações contidas no *chatbot* tem como meta esclarecer diversas dúvidas sobre os imunizantes, para que existam mais certezas e menos dúvidas sobre as vacinas,

ainda que muitas pessoas insistam em profanar mentiras sobre as mesmas.

## 5 DESENVOLVENDO UM CHATBOT PARA RESPONDER DÚVIDAS SOBRE AS VACINAS CONTRA COVID-19

Mesmo com o constante avanço das vacinações pelo Brasil e pelo mundo, não é raro encontrar pessoas com dúvidas sobre as vacinas. São diversas informações, muitas vezes incorretas ou imprecisas. Como citado anteriormente, as *fake news* e o uso de fontes pouco confiáveis acabam disseminando mais desinformação do que informação. E quem perde com isso, são as pessoas que desejam se vacinar, mas por muitas vezes ficam com medo e receio das vacinas.

### 5.1 LEVANTAMENTO DE DÚVIDAS SOBRE AS VACINAS

Com esse problema em mente, decidimos desenvolver um *chatbot*<sup>1</sup> que possui o objetivo de informar as pessoas sobre as vacinas da COVID-19. Inicialmente, coletamos através de formulários web quais são as principais dúvidas que as pessoas ainda tem sobre as vacinas.

O questionário em questão tinha como objetivo coletar os possíveis questionamentos que as pessoas ainda possuem, para servir como ponto de partida do desenvolvimento do *chatbot* proposto. Outro propósito deste questionário foi servir de fonte de dados para treinamento da NLU desenvolvida para o *chatbot*. Ou seja, usar as frases enviadas pelas pessoas para alimentar o arquivo da NLU.

Na Tabela 1, são apresentadas as perguntas mais frequentes identificadas através do questionário.

Tabela 1 – Principais dúvidas levantadas a partir de questionários.

Código	Número de Pessoas	Dúvidas
D1	8	Eficácia das vacinas (primeira e segunda dose)
D2	7	Efeitos colaterais causados pelas vacinas
D3	5	Vacinas disponíveis no Brasil
D4	5	Perguntas fora do escopo
D5	4	Como as vacinas funcionam
D6	3	Intervalo entre as doses
D7	3	Eficácia das vacinas por faixa etária
D8	2	Local de vacinação
D9	2	Eficácia das vacinas contra variantes

As dúvidas mais comuns foram sobre a eficácia das vacinas. Importante salientar que nem todos os laboratórios e farmacêuticas fornecem esses dados em seus sites ou portais oficiais. Assim, muitos dados utilizados nesse estudo foram adquiridos de estudos, artigos, levantamentos e portais de notícias conceituados e com respaldo da comunidade científica, principalmente o portal (LANCET, 2022). Também foram consultados os sites oficiais

<sup>1</sup> O *chatbot* desenvolvido está disponível na página do Github <https://github.com/letzov/tcc>.

dos fabricantes das vacinas para coletar quais as eficácias da primeira e segunda dose dos imunizantes. A engenharia de conhecimento para responder as perguntas presentes na Tabela 1 será apresentada na seção 5.3. Na seção seguinte, descrevemos como a NLU do *chatbot* foi modelada através das perguntas e suas variações coletadas através do formulário.

## 5.2 TREINAMENTO DA NLU

Solicitamos que os respondentes do formulário escrevessem suas dúvidas em formato de perguntas em linguagem natural, seguida de variações para a mesma interação, ou seja, outras perguntas que possuem o mesmo significado linguístico. As pergunta e suas respectivas variações foram utilizadas para o treinamento da unidade de NLU do *chatbot* desenvolvido. Ao todo, 39 pessoas respondem o questionário. Na tabela 1, são apresentados os resultados obtidos, onde as dúvidas dos respondentes do formulário foram agrupadas, e as 10 perguntas mais frequentes são apresentadas por assunto.

Um dos parâmetros de avaliação da qualidade e desempenho de um *chatbot* é a precisão com a qual ele reconhece o que está sendo dito pelo usuário e o que fazer com essas informações. Quando se desenvolve um *chatbot*, é preciso ter em mente aquilo que pode ser usado como entrada pelo usuário. Que tipo de mensagens e frases o usuário pode passar para o *chatbot*. Uma dos intuits do questionário foi ter uma noção de como os usuários podem realizar uma pergunta de diversas maneiras. Uma pergunta pode ser feita de maneiras distintas e ter o mesmo significado. Com isso em mente, as perguntas feitas pelas pessoas que responderam o formulário foram utilizadas no arquivo NLU do Rasa para que o treinamento fosse realizado e o *chatbot* aprendesse o padrão das perguntas. Quanto mais perguntas diferentes, mais maneiras de se extrair a intenção da pergunta e consequentemente, maior será a capacidade de reconhecimento de intenções do usuário.

Na Figura 4, estão alguns exemplos de frases que foram coletadas através do questionário e foram utilizadas para treinar o reconhecimento da intenção `eficacia_faixa_etaria` no módulo de NLU do *chatbot*.

```
- intent: eficacia_faixa_etaria
  examples: |
    - qual a eficacia das vacinas por faixa etária?
    - qual a eficiencia das vacinas por faixa etaria?
    - as vacinas tem a mesma eficacia por faixa etaria?
    - quais as eficacias por faixa etaria?
    - qual vacina funciona melhor para os mais jovens?
    - qual a vacina mais indicada para os mais velhos?
    - qual a eficacia por idade?
```

Figura 4 – Exemplo de treinamento NLU para perguntas coletadas.

Na Figura 5, é demonstrado um exemplo de outras interações que são pertinentes ao desenvolvimento do *chatbot*, mas que não foram coletas a partir do formulário, como

por exemplo o início de uma interação através de um cumprimento, reconhecido como a intenção `greet`.

```
intent: greet
examples: |
- oi
- ola
- bom dia
- boa tarde
- boa noite
- como vai?
- oie
- saudações
- oi, bom dia
- oi boa tarde
- ola boa noite
- tudo bem?
```

Figura 5 – Exemplo treinamento NLU para comunicação amigável.

### 5.3 ENGENHARIA DO CONHECIMENTO

Para responder as perguntas da Tabela 1, foi construída uma estrutura de dados no arquivo `actions.py`. A estrutura de dados em questão, consiste em um dicionário onde estão armazenados todos os dados sobre as vacinas. A vantagem de implementar essa estrutura está na possibilidade da mesma ser atualizada em tempo real. Seus dados podem ser substituídos em qualquer momento, mantendo-se a estrutura do dicionário. Isso também possibilita a adição ou exclusão de novas vacinas, e respectiva capacidade do *chatbot* responder perguntas sobre as mesmas.

Nesse estudo, foram incluídas 6 vacinas (Pfizer, Astrazeneca, Coronavac, Janssen, Sputnik e Moderna) e suas respectivas informações. Sempre que o usuário pergunta algo referente a uma vacina específica, o *chatbot* por meio da estrutura de dados desenvolvida, faz uma pesquisa no dicionário para buscar a informação correspondente. Cada vacina possui as seguintes informações na estrutura de dados:

- Eficácia contra a covid-19 após a primeira e segunda dose<sup>2</sup>;
- Eficácia contra as variantes;
- Intervalo entre cada uma das doses;
- Efeitos colaterais;
- Eficácia por faixa etária;
- Tecnologia utilizada pela vacina;

<sup>2</sup> Foram coletados dados das eficácias das vacinas em casos de internações e óbitos.

- Número de doses.

As fontes utilizadas na obtenção dos dados para montar as tabelas e formular as respostas do *chatbot* são listadas abaixo juntamente com as tabelas formuladas na estrutura de dados do *chatbot*. Cada tabela representa um conjunto de informações coletadas sobre todas as vacinas contempladas neste estudo.

1. Site oficial da fabricante da Pfizer explica a tecnologia utilizada em seu imunizante (PFIZER-BIONTECH, 2022);
2. Site da fabricante da Astrazeneca aponta as características sobre seu imunizante (NEWS, 2022);
3. Site oficial da fabricante da Janssen contém as especificações técnicas de seu imunizante (JOHNSON, 2021);
4. Site oficial do Instituto Butantan contém todas as informações sobre a vacina da Coronavac (ALIAGA; SOUZA, 2022);
5. Site oficial da fabricante da Moderna informa sobre a tecnologia utilizada em suas vacinas (MOVEMENT, 2022);
6. Site oficial da Sputnik contém as informações sobre a tecnologia utilizada em seu imunizante (HEWINGS-MARTIN, 2022);
7. Estudo realizado por cientistas da Universidade de Oxford, aponta a eficácia da vacina da Astrazeneca (primeira e segunda dose) (VOYSEY *et al.*, 2021);
8. Estudo publicado no conceituado e renomado portal Lancet sobre a eficácia da primeira e segunda dose da vacina da Pfizer (BERNAL *et al.*, 2021);
9. Site da fabricante da Janssen cita estudo publicado na publicado na New England Journal of Medicine onde foi comprovada eficácia da vacina (POWELL *et al.*, 2013);
10. Estudo liderado por desenvolvedores da fabricante da Coronavac no Brasil aponta a eficácia da segunda dose (PALACIOS *et al.*, 2021);
11. Site da fabricante da Moderna aponta eficácia da vacina (HUSSEY, 2020);
12. Estudo feito por desenvolvedores da Sputnik, publicado no portal Lancet cita a eficácia do imunizante (LOGUNOV *et al.*, 2021);

Importante salientar que, alguns governos estaduais e municipais alteraram o intervalo entre as doses das vacinas por motivos de força maior e que o intervalo de uma cidade ou estado pode ser diferente do intervalo sugerido pelos desenvolvedores das vacinas. Diferente da coleta dos dados referentes as eficacias das vacinas, a coleta dos dados

Tabela 2 – Tabela de tecnologia usada pelas vacinas da COVID-19.

Vacinas	Descrição do funcionamento e tecnologia da vacina	Fontes
Pfizer	A vacina utiliza a tecnologia chamada de mRNA ou RNA-mensageiro. Os imunizantes são produzidos por meio de replicação de sequências de RNA através de engenharia genética. O RNA mensageiro imita a proteína spike, específica do vírus Sars-CoV-2, que o ajuda a invadir as células humanas. Essa "cópia", contudo, não causa prejuízos a saúde, como o vírus, mas é o suficiente para desencadear uma reação das células do sistema imunológico, que por sua vez, cria uma defesa no organismo.	1
Astrazeneca	Usa a técnica do vetor viral não replicante. Utiliza um "vírus vivo", tal qual um adenovírus que não consegue se replicar no organismo humano ou causar malefícios a saúde. Este adenovírus também é modificado através de engenharia genética para passar a carregar dentro de si as instruções para a produção de uma proteína característica do coronavírus. Ao entrar nas células, o adenovírus faz com que elas comecem a produzir essa proteína e a exiba em sua superfície, o que é detectado pelo sistema imunológico, que age de maneira para combater o coronavírus e cria uma resposta protetora contra uma infecção	2
Janssen	A tecnologia utilizada na vacina da Janssen é baseada em vetores de adenovírus —tipo de vírus que causam o resfriado comum, mas após serem modificados para desenvolver a vacina, eles não se replicam e não causam resfriado, nem qualquer outro malefício para a saúde humana. Para realizar a produção da vacina, um pedaço da proteína "S", presente nessas espículas responsáveis pela ligação do vírus às células do corpo humano é colocado dentro do adenovírus. Quando alguém recebe a vacina composta do adenovírus não replicante, que carrega a informação genética do novo coronavírus o organismo começa um processo de defesa e produz anticorpos contra tal invasor, o que acarreta na criação de uma memória do organismo contra o coronavírus.	3
Coronavac	Feita com uma técnica conhecida como "vírus inativo". Nesta técnica, o vírus é cultivado e multiplicado e depois inativado através de uma reação química ou aquecimento. O organismo que entrar em contato com a vacina com o vírus já inativo, começa a produzir os anticorpos necessários para combater o coronavírus. As células de defesa que começam a resposta imune, encontram os vírus inativados e os capturam, ativando os linfócitos. Os linfócitos por sua vez, começam a produzir anticorpos, que entram em contato com os vírus para evitar a infecção das células.	4
Moderna	A vacina da Moderna utiliza a técnica do RNA mensageiro, ou mRNA, uma espécie de receita que imitam fragmentos do aspecto externo do coronavírus. Após a injeção da vacina, o sistema imunológico produz anticorpos contra estes fragmentos. A vacina da Moderna, contém um mRNA sintético que é capaz de codificar uma estrutura conhecida como glicoproteína de fragmento estabilizada em pré-fusão do vírus	5
Sputnik	Usa a técnica do vetor viral não replicante. A Sputnik é feita com adenovírus que são modificados em laboratório para não se repliquem após entrarem em contato com as células humanas. Os adenovírus presentes na vacina possuem instruções genéticas para a produção de uma proteína característica do coronavírus. Uma vez no organismo, os adenovírus entram nas células e fazem com que elas comecem a produzir e exibir essa proteína nas suas superfícies. Isso faz com que sistema imunológico que acione células de defesa que aprendem a combater o coronavírus.	6

Tabela 3 – Tabela de Eficácia das Vacinas contra COVID-19.

<b>Eficácia Vacinas</b>	<b>Primeira Dose</b>	<b>Segunda Dose</b>	<b>Fontes:</b>
Pfizer	61%	95%	7
Astrazeneca	76%	81%	8
Janssen	85%	Dose única	9
Coronavac	sem informação <sup>3</sup>	63%	10
Moderna	94%	94%	11
Sputnik	Sem informação <sup>4</sup>	92%	12

dos intervalos entre as doses teve mais embasamento nos sites das fabricantes do que em estudos feitos por pesquisadores e cientistas, mas tanto sites oficiais quanto artigos científicos foram utilizados para coletar estes dados.

Tabela 4 – Tabela de Intervalo entre as doses das Vacinas contra COVID-19.

Vacinas	Intervalo entre as vacinas	Fontes
Pfizer	3 Meses	13
Astrazeneca	3 Meses	14
Janssen	Dose única	15
Coronavac	21 dias	16
Moderna	28 dias	17
Sputnik	21 dias	18

13. Artigo feito por pesquisadores britânicos da universidade de Birmingham indica o intervalo ideal entre as doses da vacina da Pfizer (MOSS, 2020);
14. Site da desenvolvedora da Astrazeneca no Brasil indica o intervalo entre as doses da vacina (CRUZ, 2020);
15. Site da fabricante da Janssen aponta a necessidade de apenas uma dose da vacina para uma melhor imunização (IMMUNIZATION; DISEASES, 2020a);
16. Site do Instituto Butantan indica estudo no qual se aponta o intervalo entre as doses da vacina Coronavac (BUTANTAN, 2020b);
17. Site oficial da fabricante da Moderna indica o intervalo necessário entre as doses da vacina (IMMUNIZATION; DISEASES, 2020b);
18. Site oficial da fabricante da Sputnik aponta o intervalo entre as doses da vacina (SPUTNIK, 2020);

A única das fabricantes que disponibilizou a bula do seu respectivo imunizante foi o instituto Fiocruz, responsável pelo imunizante da Astrazeneca no Brasil. Na bula do imunizante, tem informações sobre os possíveis efeitos colaterais da vacina. Os sites dos fabricantes da Coronava e da Janssen disponibilizam informações sobre os efeitos

colaterais das suas vacinas. Estudos publicados no portal Lancet serviram de base para obter os efeitos colaterais da Sputnik, Pfizer e Moderna.

Tabela 5 – Tabela de Efeitos colaterais das Vacinas contra COVID-19.

Vacinas	Efeitos colaterais	Fontes
Pfizer	Dor no local da injeção, dor de cabeça, fadiga, dor no corpo e mal estar	19
Astrazeneca	Dor no local da injeção, dor de cabeça, fadiga, dor no corpo e mal estar	20
Janssen	Dor no local da injeção, vermelhidão, dores musculares, febre, náusea, dor no peito e falta de ar	21
Coronavac	Dor no local da injeção, febre, cansaço e calafrios	22
Moderna	Dor no local injeção, febre, vermelhidão, náusea e cansaço	23
Sputnik	Dor no local injeção, febre, vermelhidão, náusea e cansaço	24

19. Estudo publicado no portal Lancet aponta os efeitos colaterais da vacina da Pfizer (MENNI *et al.*, 2021);
20. Bula da vacina da Astrazeneca aponta possíveis efeitos colaterais do imunizante (TECNOLOGIA EM IMUNOBIOLOGICOS, 2020);
21. Página oficial da fabricante da Janssen aponta os possíveis efeitos colaterais do imunizante (THERAPEUTICS, 2020);
22. Site oficial do fabricante da Coronavac aponta os efeitos colaterais da vacina (BUTANTAN, 2020a);
23. Pesquisadores italianos publicaram artigo no portal Lancet sobre os efeitos colaterais da Moderna (OLLIARO; TORREELE; VAILLANT, 2021);
24. Estudo publicado no portal Lancet, aponta as possíveis reações ao imunizante da Sputnik (JONES; ROY, 2021);

Um artigo científico feito por pesquisadores brasileiros coletou dados sobre a eficácia das vacinas da Astrazeneca, Coronavac e Pfizer. Para as vacinas da Moderna e Janssen, o portal Observer coletou informações e dados de diversos artigos científicos sobre a eficácia desses imunizantes por faixa etária. Os desenvolvedores da Sputnik afirmaram que a vacina é igualmente eficaz em todas as faixas etárias.

25. Estudo feito por pesquisadores da Fiocruz apontam a eficácia por faixa etária das vacinas da Coronavac, Pfizer e Astrazeneca (MENEZES, 2020);
26. Levantamento de diversos estudos e artigos, feito pelo portal Observer, aponta eficácia por faixa etária dos imunizantes da Janssen e da Moderna (CAO, 2020);

Tabela 6 – Tabela de Eficácia das Vacinas contra COVID-19 por faixa etária.

Vacinas	Eficácia por Faixa Etária	Fontes
Pfizer	Acima 55 anos = 94% Entre 16 e 55 anos = 95%	25
Astrazeneca	Acima de 80 anos = 91% Entre 60 e 79 anos = 93% Entre 18 e 59 anos = 60%	25
Janssen	Acima de 59 anos = 76% Entre 18 e 59 anos = 64%	26
Coronavac	Acima de 80 anos = 68% Entre 60 e 79 anos = 79%	25
Moderna	Acima de 65 anos = 86% Entre 18 e 65 anos = 95%	26
Sputnik	Segundo desenvolvedores da Sputnik, a vacina apresenta a mesma eficiência para todas as faixas etárias	

Os fabricantes da Janssen, Coronavac e Astrazeneca publicaram em seus respectivos sites oficiais, as eficácias de seus imunizantes contra as variantes do coronavírus. Para as demais vacinas, foram coletadas em artigos científicos e estudos publicados no portal Lancet.

Tabela 7 – Tabela de Eficácia das Vacinas contra variantes da COVID-19.

Vacinas	Eficácia Contra Variantes	Fontes
Pfizer	88%	27
Astrazeneca	92%	28
Janssen	85%	29
Coronavac	Entre 69% e 78%	30
Moderna	94%	31
Sputnik	81%	32

27. Portal da fabricante no Brasil do imunizante da Astrazeneca aponta a eficácia contra variantes da vacina (AZEVEDO, 2020);
28. Estudo publicado no conceituado e renomado portal Lancet sobre a eficácia da Pfizer contra variantes (BERNAL *et al.*, 2021);
29. Site do fabricante da Janssen aponta a eficácia do imunizante contra variantes (JOHNSON; JOHNSON, 2020);
30. Site do fabricante da Coronavac aponta a eficácia do imunizante contra variantes (MOON, 2020);
31. Estudo realizado nos Estados Unidos aponta a eficácia da Moderna contra as variantes do coronavírus (TENFORDE, 2021);

32. Pesquisadores russos publicaram estudo sobre eficácia da vacina Sputnik contra variantes (BARCHUK *et al.*, 2021);

#### 5.4 DESENVOLVIMENTO DAS STORIES

Como estratégia principal para que o *chatbot* forneça as informações de acordo com o interesse do usuário, sempre que o usuário quiser saber alguma informação sobre as vacinas, o *chatbot* irá perguntar se ele gostaria de saber essa informação sobre todas as vacinas, ou sobre alguma vacina específica. Caso o usuário deseje saber sobre uma vacina em particular, o *chatbot* irá pedir para que o usuário digite o nome da vacina. Vamos usar como exemplo a *story* da Figura 6.

```
- story: quais os efeitos colaterais de uma vacina específica
  steps:
  - intent: greet
  - action: utter_how_can_I_help
  - intent: efeitos_colaterais
  - action: utter_uma_ou_todas
  - intent: saber_especifica
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    - requested_slot: nomevacina
  - active_loop: null
  - action: action_colaterais_uma_vacina
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

- story: quais os efeitos colaterais de uma vacina específica 2
  steps:
  - intent: efeitos_colaterais
  - action: utter_uma_ou_todas
  - intent: saber_especifica
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    - requested_slot: nomevacina
  - active_loop: null
  - action: action_colaterais_uma_vacina
  - action: utter_ask_user_question
  - checkpoint: check_asked_question
```

Figura 6 – Story sobre os efeitos colaterais de uma vacina específica.

Na *story* da Figura 6, é implemento um exemplo de diálogo que irá servir como base para o *chatbot* tomar suas decisões, em particular, esse exemplo implementa um diálogo para quando o usuário desejar obter informações sobre os efeitos colaterais das vacinas. A *story* da Figura 6 descreve que, após o *chatbot* reconhecer o cumprimento do usuário através da intenção `intent: greet`, ele se coloca a disposição perguntando como poderia ajudar o usuário, executando a ação `action: utter_how_can_I_help`. Então, após reconhecer que a intenção do usuário é conhecer os efeitos colaterais das vacinas, `intent: efeitos_colaterais`, ele pergunta ao usuário se ele deseja saber os efeitos colaterais de uma ou de todas as vacinas, `action: utter_uma_ou_todas`. Ao reconhecer que o usuário

deseja saber os efeitos colaterais de uma vacina específica, `intent: saber_especifica`, o *chatbot* utiliza de uma ferramenta denominada `forms`, para solicitar o nome da vacina e validar o nome na estrutura de dados disponível para ele. Após ter conhecimento do nome da vacina, e armazenado no *slot* chamado `nomevacina`, o *chatbot* é capaz de utilizar esse *slot* para fornecer a informação solicitada através da ação customizada `action: action_colaterais_uma_vacina`. Após fornecer a informação, ele ainda solicita se existe qualquer outra pergunta que o usuário gostaria de realizar, aguardando novas perguntas por parte do usuário, um cumprimento de despedida ou o fim da seção (tempo limite sem interação por parte do usuário).

```
- story: quais os efeitos colaterais das vacinas?
  steps:
  - intent: greet
  - action: utter_how_can_I_help
  - intent: efeitos_colaterais
  - action: utter_uma_ou_todas
  - intent: saber_todas
  - action: action_colaterais
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

- story: quais os efeitos colaterais das vacinas 2
  steps:
  - intent: efeitos_colaterais
  - action: utter_uma_ou_todas
  - intent: saber_todas
  - action: action_colaterais
  - action: utter_ask_user_question
  - checkpoint: check_asked_question
```

Figura 7 – *Story* sobre os efeitos colaterais de todas as vacinas.

Em contraste, a *story* da Figura 7, demonstra um exemplo que pode ser utilizado pelo *chatbot* para quando o usuário queira saber os efeitos colaterais sobre todas as vacinas disponíveis. Nessa *story*, ao reconhecer que o usuário tem a intenção de saber a informação sobre todas as vacinas, `intent: saber_todas`, ele executa a ação customizada `action_colaterais`, fornecendo a informação sobre os efeitos colaterais de todas as vacinas disponíveis na estrutura de dados utilizada.

Ambas as *stories*, mostradas nas Figuras 6 e 7, possuem o mesmo objetivo, sanar uma dúvida do usuário acerca de uma característica das vacinas, nesse caso, os efeitos colaterais, correspondente a dúvida D2 da Tabela 1. O que as difere, é que uma delas responde essa dúvida de maneira geral e a outra de maneira específica. Caso o usuário esteja interessado em saber uma informação sobre uma determinada vacina, o *chatbot* irá pedir para que o usuário digite o nome da vacina desejada. O nome da vacina é então validado através do uso de `forms`, combinado com uma método específico de validação.

Nesta etapa, é feita uma verificação do nome da vacina digitado pelo usuário. Se este nome estiver presente na lista de vacinas contida na estrutura de dados, o *chatbot* irá

fornecer as informações desejadas. Caso contrário, o usuário terá que digitar novamente o nome da vacina, até que digite um nome válido. Isso é necessário para que o *chatbot* forneça as informações corretas e não aceite qualquer nome fornecido pelo usuário, apenas os nomes disponíveis na estrutura de dados. A validação é feita através de uma ferramenta conhecida como *forms*, ferramenta essa que tem o propósito de que o *chatbot* considere válida apenas algumas informações fornecidas pelo usuário.

```
class ValidateNameForm(FormValidationAction):
    def name(self) -> Text:
        return "validate_nomevacina_form"

    def validate_nomevacina(
        self,
        slot_value: Any,
        dispatcher: CollectingDispatcher,
        tracker: Tracker,
        domain: DomainDict,
    ) -> Dict[Text, Any]:
        """Validate `nomevacina` value."""
        print(f"Nome fornecido = {slot_value}")
        if str.lower(slot_value) in vacinas:
            print(f"Nome válido = {slot_value}")
            return {"nomevacina": str.lower(slot_value)}
        else:
            print(f"Nome Inválido = {slot_value}")
            dispatcher.utter_message(text=f"Nome inválido")
            return {"nomevacina": None}
```

Figura 8 – Implementação da validação de nomes de vacinas.

O método de validação é apresentado na Figura 8, onde é feita a verificação se uma palavra enviada ao *chatbot*, pelo usuário, (neste caso, o nome da vacina) está contida em uma estrutura de dados do *chatbot*. Isto é, será feita a análise se aquela palavra em particular faz parte do domínio do chatbot. Isso ocorre por meio dos já citados *slots* de memória, que armazenam uma informação específica e comparam o valor do *slot* de memória com a estrutura de dados programada no *chatbot*. Ou seja, só serão consideradas válidas mensagens que descrevam corretamente os nomes das vacinas contempladas no *chatbot*. Essa ferramenta é essencial para que o *chatbot* e o usuário tenham uma conversa coesa e coerente. O *forms* é essencial para que as mensagens do usuário sejam devidamente compreendidas pelo *chatbot*, isto é, em um dado contexto, o *chatbot* precisa receber uma mensagem a qual ele consiga interpretar corretamente.

Outra ferramenta utilizada na elaboração deste *chatbot*, são os checkpoints. Os checkpoints são implementados nas *stories* com o objetivo de fazer o *chatbot* estar sempre ativo, isto é, independente do momento ou do contexto, se o usuário fizer uma pergunta, o *chatbot* deve responder a mensagem do usuário. Foram implementados checkpoints com a ação *restart*, uma ação padrão do *rasa* que reinicia o *chatbot*, como pode ser observado na Figura 9. Por exemplo, o usuário está interessado em descobrir a eficácia de uma vacina específica. Nesta situação, o *chatbot* irá executar uma *storie* completa, começando pela

saudação do usuário. Após fornecer os dados, o *chatbot* pergunta se o usuário deseja saber algo a mais. Caso a resposta seja sim, o *chatbot* será reiniciado. O usuário então poderá perguntar algo para o *chatbot* que por sua vez, irá responder.

```

- story: usuario continua
  steps:
  - checkpoint: check_asked_question
  - intent: positivo
  - action: action_restart

- story: usuario para
  steps:
  - checkpoint: check_asked_question
  - intent: negativo
  - action: utter_fim
  - checkpoint: check_flow_finished

- story: fim fluxo
  steps:
  - checkpoint: check_flow_finished
  - intent: fim_chatbot
  - action: utter_goodbye

```

Figura 9 – Implementação de checkpoints.

Nas Figuras 10 e 11 são apresentadas as *stories* para responder dúvidas sobre a eficácia das vacinas. Essas *stories* tem uma peculiaridade em relação as demais *stories*. Existem duas *stories* distintas, uma para cada dose das vacinas. Isto implica que caso o usuário deseje saber a eficácia sobre uma dose específica de um imunizante em particular, o *chatbot* irá responder adequadamente.

```

- story: qual a eficacia de uma vacina (primeira dose)
  steps:
  - intent: eficacia_imediata
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    - requested_slot: nomevacina
  - active_loop: null
  - action: utter_primeira_ou_segunda
  - intent: primeira_dose
  - action: action_eficacia_primeira_dose
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

```

Figura 10 – Storie sobre a eficácia da primeira dose

As Figuras 12 e 13, são apresentadas as *stories* que respondem questões sobre o intervalo entre as doses. Elas possuem roteiro similar as demais *stories*, considerando casos em que o usuário deseja saber intervalos de uma vacina específica ou de todas

```

- story: qual a eficacia de uma vacina (segunda dose)
  steps:
  - intent: eficacia_imediata
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    | - requested_slot: nomevacina
  - active_loop: null
  - action: utter_primeira_ou_segunda
  - intent: segunda_dose
  - action: action_eficacia_segunda_dose
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

```

Figura 11 – Stories sobre a eficácia da segunda dose.

elas. Em particular, a *story* da Figura 13, permite que o *chatbot* identifique o nome da vacina específica diretamente na pergunta realizada pelo usuário, fornecendo a informação diretamente à ele.

```

- story: quais os intervalos entre as doses das vacinas
  steps:
  - intent: intervalo_vacinas
  - action: utter_uma_ou_todas
  - intent: saber_todas
  - action: action_intervalo_todas_doses
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

- story: qual o intervalo entre as doses de uma vacina específica
  steps:
  - intent: intervalo_vacinas
  - action: utter_uma_ou_todas
  - intent: saber_especifica
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    | - requested_slot: nomevacina
  - active_loop: null
  - action: action_intervalo_doses
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

```

Figura 12 – Stories sobre o intervalo entre as doses das vacinas.

Outra dúvida pertinente levantada pelo questionário foi sobre o interesse em que as pessoas possuem em saber sobre a eficácia das vacinas considerando as diferentes faixas etárias. Na Figura 14 e 15, demonstramos as *stories* desenvolvidas para responder dúvidas sobre a eficácia das vacinas por faixa etária. Cabe salientar que não há a coleta da idade do usuário, apenas são fornecidas as informações referente a eficácia das vacinas por faixa

```

- story: qual o intervalo imediato? 2
  steps:
  - intent: intervalo_imediato
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    - requested_slot: nomevacina
  - active_loop: null
  - action: action_intervalo_doses
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

```

Figura 13 – Stories sobre o intervalo entre as doses das vacinas imediato.

etária, conforme o interesse do usuário em saber essa informação sobre todas as vacinas ou sobre uma vacina específica.

```

- story: qual a eficácia de uma vacina por faixa etária
  steps:
  - intent: eficacia_faixa_etaria
  - action: utter_uma_ou_todas
  - intent: saber_especifica
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    - requested_slot: nomevacina
  - active_loop: null
  - action: action_faixa_etaria_individual
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

- story: quais as eficácias por faixa etária das vacinas
  steps:
  - intent: eficacia_faixa_etaria
  - action: utter_uma_ou_todas
  - intent: saber_todas
  - action: action_faixa_etaria_todas
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

```

Figura 14 – Stories sobre eficácia das vacinas por faixa etária.

Para responder as dúvidas sobre a eficácias dos imunizantes em relação as variantes, foram desenvolvidas as *stories* da Figura 16. Também foram desenvolvidas *stories* que prevem o cenário onde o usuário deseja saber a informação sobre uma vacina específica imediatamente, logo na primeira interação. Seu funcionamento é semelhante as demais *stories*, onde ela prevê a opção do usuário desejar saber essa informação sobre uma única

```

- story: qual a faixa etaria imediata ?
steps:
- intent: faixa_etaria_imediata
- action: nomevacina_form
- active_loop: nomevacina_form
- slot_was_set:
  - requested_slot: nomevacina
- active_loop: null
- action: action_faixa_etaria_individual
- action: utter_ask_user_question
- checkpoint: check_asked_question

```

Figura 15 – Stories sobre eficácia das vacinas por faixa etária imediata.

vacina ou sobre todas elas.

```

- story: qual a eficacia de todas as vacinas contra as variantes
steps:
- intent: variantes_vacinas
- action: utter_uma_ou_todas
- intent: saber_todas
- action: action_inform_eficienciatodas_variantes
- action: utter_ask_user_question
- checkpoint: check_asked_question

- story: qual eficacia contra as variantes de uma vacina específica
steps:
- intent: variantes_vacinas
- action: utter_uma_ou_todas
- intent: saber_especifica
- action: nomevacina_form
- active_loop: nomevacina_form
- slot_was_set:
  - requested_slot: nomevacina
- active_loop: null
- action: action_inform_eficacia_uma_variantes
- action: utter_ask_user_question
- checkpoint: check_asked_question

```

Figura 16 – Stories eficácia contra variantes

Para responder as dúvidas sobre as tecnologias utilizadas nos imunizantes, foram desenvolvidas as *stories* das Figuras 17 e 18. A estratégia utilizada nessas *stories* foi similar as demais, onde o usuário pode perguntar diretamente a informação da tecnologia de uma vacina específica, e o *chatbot* irá utilizar a *story* da Figura 18 para fornecer a informação diretamente, ou ainda ele pode realizar um questionamento quanto a característica sem mencionar a vacina, e o *chatbot* irá prosseguir perguntando se ele deseja saber a informação sobre uma vacina específica ou sobre todas as vacinas, seguindo as *stories* da Figura 17.

```

- story: como uma vacina funciona
  steps:
  - intent: como_funcionam
  - action: utter_uma_ou_todas
  - intent: saber_especifica
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    | - requested_slot: nomevacina
  - active_loop: null
  - action: action_tecnologia_especifica
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

- story: como funcionam todas |
  steps:
  - intent: como_funcionam
  - action: utter_uma_ou_todas
  - intent: saber_todas
  - action: action_tecnologia_todas
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

```

Figura 17 – Stories Tecnologia vacinas

```

- story: como funciona uma vacina específica?
  steps:
  - intent: como_funcionam_imediata
  - action: nomevacina_form
  - active_loop: nomevacina_form
  - slot_was_set:
    | - requested_slot: nomevacina
  - active_loop: null
  - action: action_tecnologia_especifica
  - action: utter_ask_user_question
  - checkpoint: check_asked_question

```

Figura 18 – Stories Tecnologia vacinas imediata

Concluindo, as *stories* apresentadas nessa seção implementam as diferentes interações que o *chatbot* pode utilizar para responder as perguntas que foram levantadas através de uma consulta pública sobre as dúvidas que permeiam as vacinas contra COVID-19, sumarizadas na Tabela 1 da Seção 5.1. Todas as *stories* utilizam da estrutura de dados modelada a partir dos levantamentos realizados e apresentados na Seção 5.3, onde as informações pertinentes sobre cada uma das vacinas é consultada, e uma resposta em linguagem natural é elaborada. Essas consultas e elaboração de respostas são implemen-

tadas através de ações customizáveis, indicadas em cada *story* pela interação `action:`  
`action_<nome da ação customizável desenvolvida>`.

## 6 AVALIAÇÃO DO CHATBOT DESENVOLVIDO

Após o término do processo de desenvolvimento do chatbot, realizamos uma avaliação empírica utilizando a ferramenta de testes fornecida pelo próprio Rasa, a fim de verificar a precisão no reconhecimento das intenções do usuário e se as *stories* desenvolvidas eram de fato executadas de maneira adequada.

Para realização das análises apresentadas nessa seção, foram utilizadas ferramentas de avaliação disponíveis juntamente com o *framework* do rasa. A avaliação nos permite uma visualização gráfica e numérica da precisão e acurácia do *chatbot*.

A avaliação do *chatbot* compreende em 2 principais aspectos: (i) a avaliação da precisão do reconhecimento de intenções e entidades pela NLU, e (ii) a avaliação das *stories* desenvolvidas através do rasa core.

### 6.1 AVALIAÇÃO DA NLU

Para realizar a avaliação da NLU do *chatbot*, o processo compreende a realização das 4 etapas a seguir:

1. Criar uma divisão das estradas presentes no arquivo nlu.yml. Esta divisão por padrão é feita de maneira que 80 por cento do arquivo é selecionado para treinamento, enquanto que os 20 por cento restante são utilizados para testes;
2. Excluir os dados selecionados para teste do arquivo utilizado para treinamento;
3. Realiza o treinamento da NLU com o arquivo remanescente;
4. Avaliação do modelo gerado utilizando os dados de teste.

Após a realização desta etapa, são obtidos gráficos gerados pela ferramenta de testes do Rasa. Neste gráfico, é possível observar a precisão com a qual as intenções são reconhecidas pelo *chatbot*. Cada linha e coluna representa uma intenção. No eixo vertical estão presentes as intenções que deveriam ser previstas nos testes, e no eixo horizontal, as intenções que foram efetivamente previstas nos testes executados. Um número indicando a quantidade de vezes que uma intenção foi devidamente prevista pelo *chatbot* será mostrado na diagonal principal da matriz.

Na matriz da Figura 20 são apresentados os resultados obtidos. É possível observar que a intenção “saber-todas” não possui entradas de treinamento e teste. Isso ocorre pois não foram feitos exemplos para esta intenção no arquivo da NLU, pois tal intenção foi usada apenas como função de escolha no *chatbot*. Ou seja, o usuário deseja saber uma informação sobre as vacinas e tem a opção de saber sobre uma específica ou sobre todas. O resultado dessa escolha não é por meio de interação textual em linguagem natural e sim por meio de botões de escolha e portanto não necessita de exemplos na NLU. Para a intenção

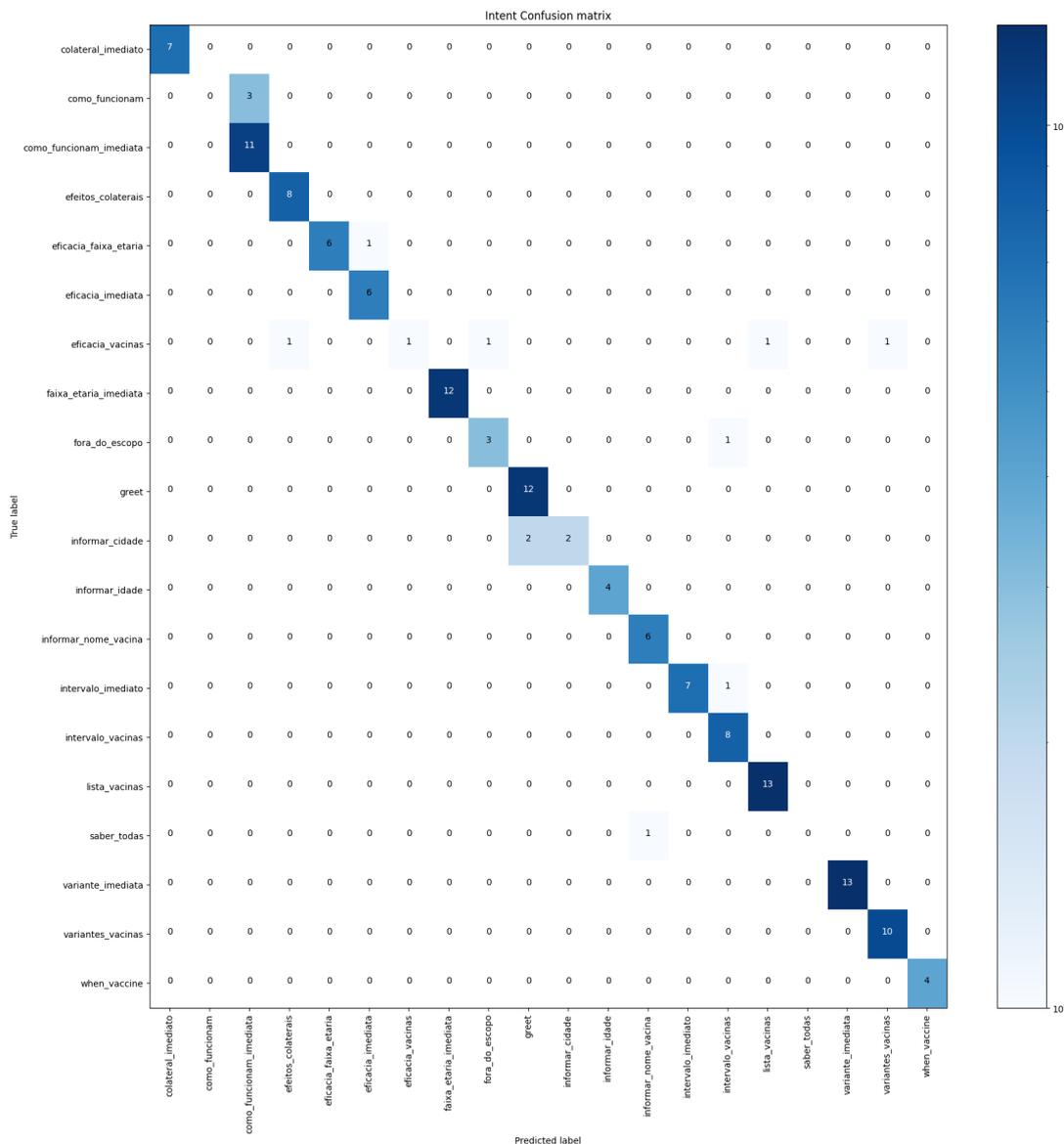


Figura 19 – Matriz de confusão com resultados dos testes da NLU.

"informar-cidade", em dois casos foi prevista equivocadamente a intenção "greet", isto é, a intenção de saudação do *chatbot*. Isto ocorre pois para nomes de cidades, existem diversos exemplos, alguns podem ter sido interpretados como saudações. Em trabalhos futuros, pode-se acrescentar um arquivo externo contendo todos os municípios do Brasil, e fazer com que o *chatbot* considere como válidas apenas entradas pertencentes a este arquivo. Por fim, a intenção "eficacia-vacinas" foi acertadamente prevista, mas outras quatro vezes prevista de maneira equivocada. Como citado anteriormente, existem diversas maneiras de se enviar uma mesma mensagem. Como o *chatbot* tem de responder diversas mensagens sobre características das vacinas, algumas muito parecidas (ou seja, o *chatbot* previu erroneamente "eficacia-vacina" com "eficacia-imediata"), é possível que algumas intenções tenham sido previstas de maneira errada. É possível solucionar este problema aumentando o número de exemplos para cada intenção no módulo de NLU. As demais intenções

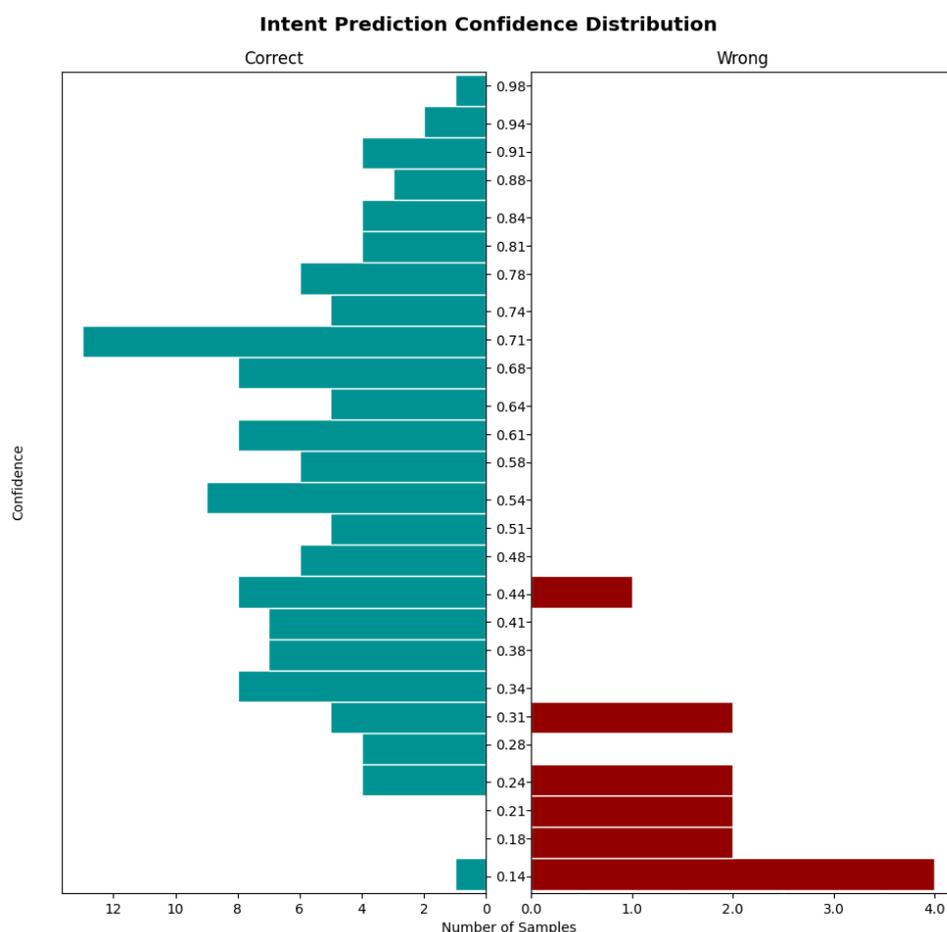


Figura 20 – Histograma de distribuição de confiança no reconhecimento de intenções.

apresentaram um resultado satisfatório.

No histograma da Figura 20, é apresentado a quantidade de erros e acertos em relação a confiança na predição. Como citado anteriormente, quanto mais exemplos de treinamento contidos no *chatbot*, maior a chance do mesmo prever corretamente a intenção do usuário. As barras verdes representam o número de acertos enquanto as vermelhas o número de erros nas previsões das intenções. Os números presentes no eixo vertical representam um índice de 0 a 1, onde quanto mais perto de 1, maior a precisão de reconhecimento da intenção, e os números do eixo horizontal representam a quantidade de exemplos classificados corretamente ou erroneamente dependendo da confiança de predição. Para eventuais trabalhos futuros, é possível observar o índice de precisão na qual as intenções são equivocadamente previstas e estabelecer uma métrica onde, se o reconhecimento da intenção do usuário tiver uma precisão abaixo deste índice, o *chatbot* pode solicitar para o usuário digitar sua mensagem novamente, sugerindo que o mesmo digite a pergunta de uma forma diferente com o objetivo de entendê-lo melhor. Isso, possivelmente, implicaria em um melhor reconhecimento das intenções do usuário, pois mensagens com baixo índice de confiança são descartadas e conseqüentemente, o chatbot iria responder apenas intenções identificadas com uma maior precisão.

```
- story: qual a eficacia de todas as vacinas
steps:
- user: |
  | qual a eficacia das vacinas?
  intent: eficacia_vacinas
- action: utter_uma_ou_todas
- user: |
  | saber todas
- action: action_inform_eficienciatodas
- action: utter_ask_user_question
- checkpoint: check_asked_question
```

Figura 21 – Exemplo de *story* de teste desenvolvida.

Além da matriz de confusão e do histograma apresentados, a ferramenta de testes também fornece as informações de f1-score<sup>1</sup>, acurácia, e precisão, tanto para classificação de intenções, como para extração de entidades.

É possível constatar que a maioria dos valores dos testes são satisfatórios. Como resultado, pode-se esperar que chatbot seja capaz de prever adequadamente as intenções do usuário. O que será constatado na seção seguinte com a avaliação das *stories* desenvolvidas, que possuem dependência direta do reconhecimento adequado das intenções e entidades extraídas a partir das interações do usuário.

## 6.2 AVALIAÇÃO DAS *STORIES*

Para analisar a eficácia e precisão do *chatbot* com relação as *stories*, foram desenvolvidas *stories* de teste (por exemplo, a *story* de teste da Figura 21), nas quais são simuladas conversas entre o *chatbot* e o usuário. Os testes especificam, dada a mensagem recebida, quais intenções devem ser reconhecidas e quais ações o *chatbot* deveria realizar.

O número de testes pode ser alterado de acordo com o desejo do desenvolvedor. Por padrão, são realizados 200 testes. Uma vez programadas, as *stories* de teste são executadas pela ferramenta de teste do Rasa de maneira que as entradas, isto é, as mensagens enviadas pelo usuário são as mesmas descritas nas *stories* de teste juntamente com outras mensagens similares descritas na NLU. Após a realização dos testes, o rasa disponibiliza alguns gráficos que ajudam a perceber o quão preciso foram as interações. Tais gráficos mostram quantas ações foram previstas corretamente de acordo com o que estava nas *stories* de teste. Foram desenvolvidas uma *story* de teste para cada *story* contida no projeto do chatbot. Os resultados dos testes podem ser observados na Figura 22.

Na Figura 22, os resultados podem ser vistos em forma de matriz. Estão contidos no eixo horizontal as ações que as *stories* deveriam prever após serem executadas, enquanto

<sup>1</sup> O f1-score representa uma métrica de avaliação que mede a precisão de testes, ele varia de 0 a 1, onde quanto mais próximo de 1, mais preciso foi o resultado.

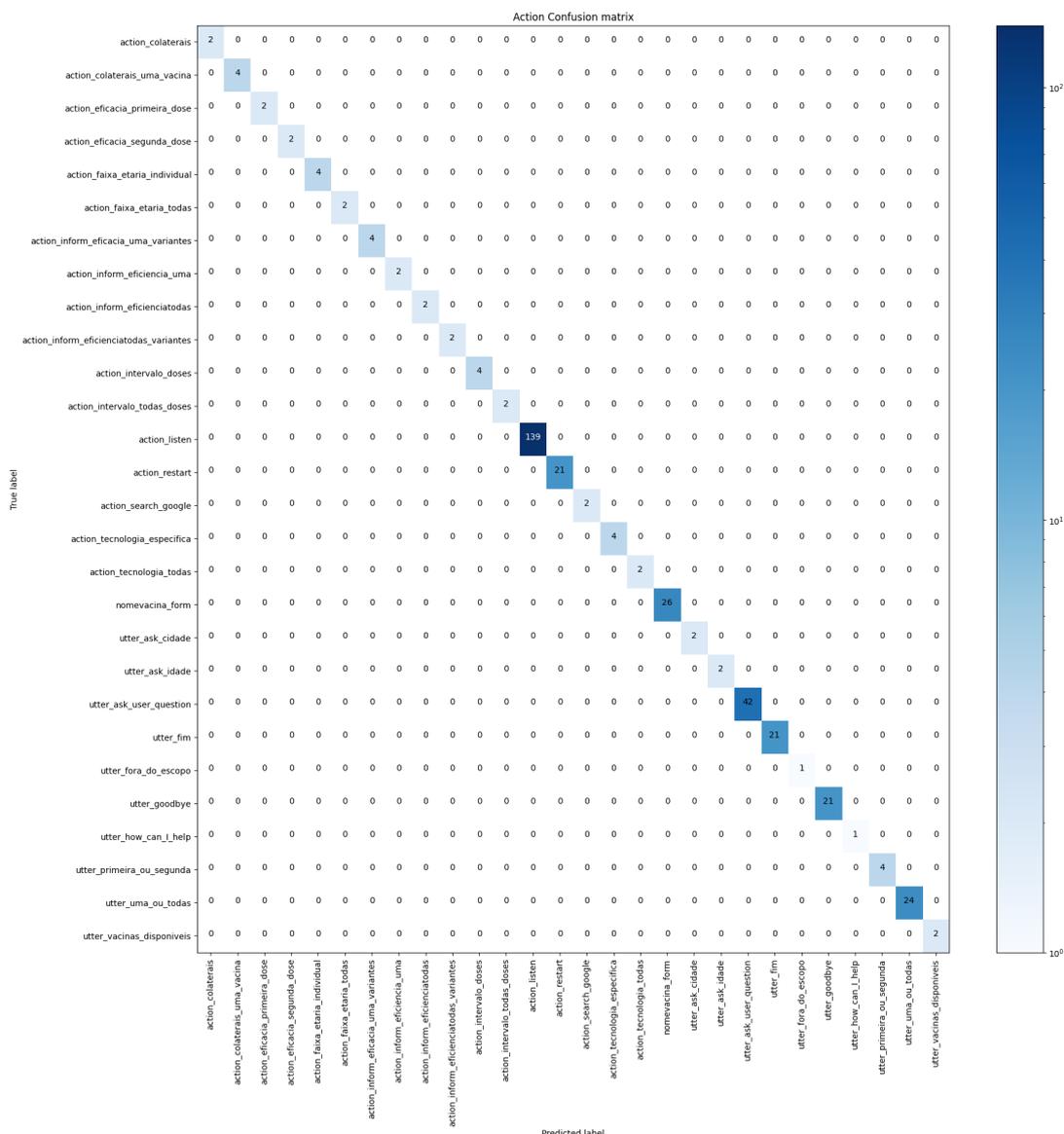


Figura 22 – Matriz de confusão com resultado dos testes de stories.

que no eixo vertical estão as ações que foram previstas pelo *chatbot* na fase de testes. Isto significa que para cada ação prevista corretamente pelo *chatbot*, é acrescentado uma contabilização na diagonal principal. Este número representa o número de vezes que tal ação foi identificada durante a fase de testes e ao aparecer na diagonal principal, implica que ela foi reconhecida corretamente. Caso alguma ação tivesse sido erroneamente prevista, o número de vezes que tal ação foi identificada de maneira equivocada seria alocado fora da diagonal principal da matriz.

Observa-se que as classificações ocorrem apenas na diagonal principal da matriz, o que implica em uma precisão satisfatória do *chatbot*. A respeito da precisão do *chatbot*, podemos notar que o mesmo atingiu um grau satisfatório, evidenciado principalmente na Figura 22, mesmo com algumas classificações equivocadas, o *chatbot* é capaz de conduzir corretamente as interações, respondendo de maneira satisfatória.

## 7 CONCLUSÃO E TRABALHOS FUTUROS

No atual contexto no qual o mundo está inserido, a busca por informações completas e verídicas sobre diversos temas tem se tornado um dos pilares da sociedade. Independente do tema, surgem diversas informações em poucos segundos na nossa frente, graças ao advento da internet e das redes sociais. Contudo, nem toda a informação que consumimos é verídica. As *fake news* infelizmente são parte do cotidiano. A pandemia da covid-19 fez com que a necessidade por informações verídicas fosse ainda mais urgente. Uma informação falsa ou mal intencionada e em poucos segundos uma mentira pode ser tratada como se fosse algo verídico. Para prover uma fonte de informações confiáveis, foi desenvolvido um *chatbot* capaz de sanar dúvidas sobre as vacinas da covid-19.

A capacidade de atualização das informações contidas no *chatbot*, a praticidade no manuseio da tecnologia e o fácil acesso ao mesmo estão entre os principais motivos pelos quais essa tecnologia foi empregada nesse projeto.

Após o término de todo o processo de desenvolvimento do *chatbot*, foi possível constatar que o *chatbot* possui uma eficiência satisfatória ao receber as entradas do usuário, o que implica que o objetivo principal de desenvolver um *chatbot* capaz de responder dúvidas sobre as vacinas da covid-19, foi alcançado. Após uma série de testes, pode-se notar empiricamente a precisão de reconhecimento do *chatbot* ao receber diversas entradas fornecidas pelo usuário. Também destaca-se a facilidade com a qual o *chatbot* pode ser atualizado a medida que o contexto das vacinas e da pandemia da covid-19 avançam, considerando a estrutura de dados modelada. Também é válido destacar que os objetivos e metas iniciais foram todos alcançados. O *chatbot* é capaz reconhecer e interpretar todas as dúvidas e mensagens para qual foi desenvolvido para responder. Por fim, é pertinente afirmar que o *chatbot* é uma poderosa ferramenta de conhecimento, uma vez que possui alto grau de acessibilidade, clareza e transparência com o usuário, características ideais para sanar as mais diversas dúvidas que o atual contexto da sociedade oferece.

### 7.1 TRABALHOS FUTUROS

Em um eventual cenário de novas vacinas, doses e variantes, são necessários poucos ajustes para que o *chatbot* esteja sempre atualizado com o meio ao qual está inserido.

A eficácia das doses de reforço, a eficácia contra novas variantes e eventuais novas fabricantes de vacina ou novos dados sobre as vacinas já existentes e podem ser adicionadas ao *chatbot* em trabalhos futuros.

Também é válido considerar tornar o *chatbot* público de alguma forma, seja disponibilizá-lo em um site ou em aplicativos de mensagem, como Telegram e Whatsapp.

Porém, antes de disponibilizar o chatbot, seria interessante realizar testes com usuários para ter uma noção se o *chatbot* está cumprindo as funções as quais foi programado para executar bem como o nível de satisfação do usuário.

## REFERÊNCIAS

- ADAMOPOULOU, Eleni; MOUSSIADES, Lefteris. Chatbots: History, technology, and applications. **Machine Learning with Applications**, Elsevier, v. 2, p. 100006, 2020.
- ALERT, WHO Health. **WHO launches a chatbot on Facebook Messenger to combat COVID-19 misinformation**. [S.l.]: WHO Health Alert, 2022. Disponível em: <https://www.who.int/news-room/feature-stories/detail/who-launches-a-chatbot-powered-facebook-messenger-to-combat-covid-19-misinformation>. Acesso em: 13 de fevereiro 2022.
- ALIAGA, Laura Isabel Torres; SOUZA, Patrícia Rodrigues Rezende de. Vacina e suas tecnologias. **ANALECTA-Centro Universitário Academia**, v. 7, n. 2, 2022.
- ALLEN, James F. Natural language processing. *In*: **ENCYCLOPEDIA of computer science**. [S.l.: s.n.], 2003. P. 1218–1222.
- AZEVEDO, Cristina. <https://portal.fiocruz.br/en/news/astrazeneca-vaccine-92-effective-against-hospitalizations-due-delta-variant>. Edição: Portal Fio Cruz. [S.l.: s.n.], 2020.
- BARCHUK, Anton *et al.* Vaccine Effectiveness against Referral to Hospital and Severe Lung Injury Associated with COVID-19: A Population-based Case-control Study in St. Petersburg, Russia. **medRxiv**, Cold Spring Harbor Laboratory Press, 2021.
- BARRETT, Philip *et al.* **After-Effects of the COVID-19 Pandemic: Prospects for Medium-Term Economic Damage**. [S.l.], 2021.
- BERNAL, Jamie Lopez *et al.* Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. **bmj**, British Medical Journal Publishing Group, v. 373, 2021.
- BUTANTAN, Instituto. <https://butantan.gov.br/covid/butantan-tira-duvida/tira-duvida-noticias/coronavac-possui-alto-perfil-de-seguranca-e-baixa-ocorrencia-de-reacoes-adversas-conheca-as-mais-comuns>. Edição: Instituto Butantan. [S.l.: s.n.], 2020.
- \_\_\_\_\_. <https://butantan.gov.br/noticias/eficacia-global-da-coronavac-pode-chegar-a-623-com-intervalo-entre-doses-igual-ou-superior-a-21-dias>. Edição: Instituto Butantan. [S.l.: s.n.], 2020.
- CAO, Sissi. <https://observer.com/2021/04/covid19-vaccine-efficacy-rate-age-group-pfizer-moderna-jj-biontech/>. Edição: Observer. [S.l.: s.n.], 2020.

CRUZ, Fundação Oswaldo. <https://portal.fiocruz.br/noticia/covid-19-fiocruz-esclarece-sobre-intervalo-de-doses-da-vacina>. Edição: Fundação Oswaldo Cruz. [S.l.: s.n.], 2020.

FINATTO, Maria José Bocorny; LOPES, Lucelene; SILVA, Alena Ciulla. Processamento de Linguagem Natural, Linguística de Corpus e Estudos Linguísticos: uma parceria bem-sucedida. **Domínios de linguagem**. Uberlândia, MG. Vol. 9, n. 5 (dez. 2015), p. 41–59, 2015.

HEWINGS-MARTIN, Yella. **How does the Sputnik V COVID-19 vaccine work?** [S.l.: s.n.], 2022. Disponível em: <https://www.medicalnewstoday.com/articles/how-does-the-sputnik-v-covid-19-vaccine-work>. Acesso em: 13 de fevereiro 2022.

HUSSEY, Colleen. **IE9 for Windows Phone 7: Adobe Flash, demos and development**. Edição: Lavina Talukdar. [S.l.: s.n.], 2020. Disponível em: <https://investors.modernatx.com/news-releases/news-release-details/moderna-announces-primary-efficacy-analysis-phase-3-cove-study/>.

IMMUNIZATION, National Center for; DISEASES, Respiratory. <https://www.cdc.gov/vaccines/covid-19/info-by-product/janssen/janssen-faqs.html>. Edição: National Center for Immunization e Respiratory Diseases. [S.l.: s.n.], 2020.

\_\_\_\_\_. <https://www.cdc.gov/vaccines/covid-19/info-by-product/moderna/index.html>. Edição: National Center for Immunization e Respiratory Diseases. [S.l.: s.n.], 2020.

INÁCIO, Luiz Gustavo Kempe; CANATO, Robson Leandro Carvalho. PROPOSTA DE DESENVOLVIMENTO DE UM CHATBOT UTILIZANDO A PLATAFORMA DIALOGFLOW. **Interciência & Sociedade**, v. 5, n. 2, p. 564–582, 2020.

JOHNSON, J. **Johnson & Johnson COVID-19 Vaccine Authorized by US FDA For Emergency Use-First Single-Shot Vaccine in Fight Against Global Pandemic**. [S.l.]: Johnson & Johnson New Brunswick, NJ, USA, 2021.

JOHNSON; JOHNSON. <https://www.jnj.com/positive-new-data-for-johnson-johnson-single-shot-covid-19-vaccine-on-activity-against-delta-variant-and-long-lasting-durability-of-response>. Edição: Johnson e Johnson. [S.l.: s.n.], 2020.

JONES, Ian; ROY, Polly. Sputnik V COVID-19 vaccine candidate appears safe and effective. **The Lancet**, Elsevier, v. 397, n. 10275, p. 642–643, 2021.

KANNAN, PV; BERNOFF, Josh. Does your company really need a chatbot? **Harvard Business Review**, Digital articles, p. 2–6, 2019.

KANNAN, Subbu; GURUSAMY, Vairaprakash *et al.* Preprocessing techniques for text mining. **International Journal of Computer Science & Communication Networks**, v. 5, n. 1, p. 7–16, 2014.

KUYVEN, Neiva Larisane *et al.* Chatbots na educação: uma Revisão Sistemática da Literatura. **RENOTE**, v. 16, n. 1, 2018.

LANCET. **The Lancet**. [S.l.: s.n.], 2022. Last accessed 22 February 2022. Disponível em: <https://www.thelancet.com/>.

LOGUNOV, Denis Y *et al.* Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. **The Lancet**, Elsevier, v. 397, n. 10275, p. 671–681, 2021.

LUGLI, Verônica Adelaide; LUCCA FILHO, João de. O USO DO CHATBOT PARA A EXCELÊNCIA EM ATENDIMENTO. **Revista Interface Tecnológica**, v. 17, n. 1, p. 205–218, 2020.

MENEZES, Maíra.

[https://agencia.fiocruz.br/sites/agencia.fiocruz.br/files/u34/nt\\_efetividade\\_vacinas.pdf](https://agencia.fiocruz.br/sites/agencia.fiocruz.br/files/u34/nt_efetividade_vacinas.pdf)  
Edição: Fundação Oswaldo Cruz. [S.l.: s.n.], 2020.

MENNI, Cristina *et al.* Vaccine side-effects and SARS-CoV-2 infection after vaccination in users of the COVID Symptom Study app in the UK: a prospective observational study. **The Lancet Infectious Diseases**, Elsevier, 2021.

MOON, Peter. <https://butantan.gov.br/noticias/estudo-mostra-que-coronavac-e-eficaz-contra-casos-graves-de-covid-19-causados-pela-variante-delta>. Edição: Instituto Butantan. [S.l.: s.n.], 2020.

MOSS, Paul. <https://www.birmingham.ac.uk/news/latest/2021/05/covid-pfizer-vaccination-interval-antibody-response.aspx>. Edição: Emma McKinney. [S.l.: s.n.], 2020.

MOURA, Marilís Pereira Lima *et al.* Prospecção de Tecnologias em Chatbots. **Cadernos de Prospecção**, v. 12, n. 3, p. 550–550, 2019.

MOVEMENT, Patient Safety (Ed.). **COVID-19 Vaccine Information**. [S.l.: s.n.], 2022. Disponível em: [https://patientsafetymovement.org/helpful-coronavirus-covid-19-resources/covid-19-vaccine-information/?gclid=CjwKCAiAo40QBhBBEiwA5KWu\\_7ikhRIwPw1s5B5mxjg2SAQRZ9MwdCiJYxfSWR6pULv40Yj1qoP-VxoCZEAQAvD\\_BwE](https://patientsafetymovement.org/helpful-coronavirus-covid-19-resources/covid-19-vaccine-information/?gclid=CjwKCAiAo40QBhBBEiwA5KWu_7ikhRIwPw1s5B5mxjg2SAQRZ9MwdCiJYxfSWR6pULv40Yj1qoP-VxoCZEAQAvD_BwE) Acesso em: 13 de fevereiro 2022.

NAILI, Marwa; CHAIBI, Anja Habacha; GHEZALA, Henda Hajjami Ben. Comparative study of word embedding methods in topic segmentation. **Procedia computer science**, Elsevier, v. 112, p. 340–349, 2017.

NEWS, BBC. **Covid: How does the Oxford-AstraZeneca vaccine work?** [S.l.: s.n.], 2022. Disponível em: <https://www.bbc.com/news/health-55302595>. Acesso em: 13 de fevereiro 2022.

OLLIARO, Piero; TORREELE, Els; VAILLANT, Michel. COVID-19 vaccine efficacy and effectiveness—the elephant (not) in the room. **The Lancet Microbe**, Elsevier, 2021.

ORGANIZATION, World Health. **WHO Coronavirus (COVID-19) Dashboard**. [S.l.: s.n.], 2022. Last accessed 22 February 2022. Disponível em: <https://covid19.who.int/>.

PALACIOS, Ricardo *et al.* Efficacy and safety of a COVID-19 inactivated vaccine in healthcare professionals in Brazil: the PROFISCOV study, 2021.

PEREIRA, Gabrielle; PINHEIRO, Marco Antônio. Conversando com robôs: O uso de chatbots na comunicação de marcas no Facebook Messenger1, 2018.

PFIZER-BIONTECH. **Centers for Disease Control and Prevention**. [S.l.: s.n.], 2022. Disponível em: <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/Pfizer-BioNTech.html>. Acesso em: 13 de fevereiro 2022.

PIZZAHUT. **ChatbotGuide**. [S.l.: s.n.], 2017. Last accessed 22 February 2022. Disponível em: <https://www.chatbotguide.org/pizzahut-bot>.

POWELL, Jerry S *et al.* Phase 3 study of recombinant factor IX Fc fusion protein in hemophilia B. **New England Journal of Medicine**, Mass Medical Soc, v. 369, n. 24, p. 2313–2323, 2013.

ROSSATO, Danúbia Miorando *et al.* Engajamento de trabalhadores na implementação de chatbot para atendimento aos usuários de uma universidade. Universidade de Passo Fundo, 2020.

\_\_\_\_\_. Engajamento de trabalhadores na implementação de chatbot para atendimento aos usuários de uma universidade. Universidade de Passo Fundo, 2020.

SHARMA, Rakesh Kumar; JOSHI, Manoj. An analytical study and review of open source chatbot framework, rasa. **International Journal of Engineering Research and**, v. 9, n. 06, 2020.

SIENKIEWICZ, Agnieszka. **11 Chatbot Statistics and Trends You Need to Know in 2021**. [S.l.]: Tidio, 2022. Disponível em:

<https://www.tidio.com/blog/chatbot-statistics/>. Acesso em: 13 de fevereiro 2022.

SOUZA, Gustavo Oliveira de; RIBEIRO, Victor M; BRETERNITZ, Vilvaldo José. Um estudo e implementação de chatbots utilizando uma arquitetura orientada a serviços. Universidade Presbiteriana Mackenzie, 2019.

SPUTNIK. [https://rdif.ru/Eng\\_fullNews/6413/](https://rdif.ru/Eng_fullNews/6413/). Edição: Sputnik. [S.l.: s.n.], 2020.

TECNOLOGIA EM IMUNOBIOLOGICOS, Instituto de. <https://www.bio.fiocruz.br/images/bula-vacina-covid-19-recombinante-vp-002-27-01-2021.pdf>. Edição: Fundação Oswaldo Cruz. [S.l.: s.n.], 2020.

TENFORDE, Mark W. Effectiveness of Pfizer-BioNTech and Moderna Vaccines Against COVID-19 Among Hospitalized Adults Aged less than 65 Years—United States, January–March 2021. **MMWR. Morbidity and mortality weekly report**, v. 70, 2021.

THERAPEUTICS, Janssen. <https://www.janssencovid19vaccine.com/hcp/frequently-asked-questions.html>. Edição: Janssen. [S.l.: s.n.], 2020.

TUCUNDUVA, RODRIGO. **Exemplos de chatbot: 6 empresas que usam a tecnologia**. [S.l.: s.n.], 2022. Last accessed 22 February 2022.

VOYSEY, Merryn *et al.* Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials. **The Lancet**, Elsevier, v. 397, n. 10277, p. 881–891, 2021.