UNIVERSIDADE DE SÃO PAULO
ESCOLA DE COMUNICAÇÃO E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

RAFAEL JOÃO DA SILVA

Batendo um papo com a Informação: O uso dos *chatbots* para a recuperação da informação e a contribuição da Ciência da Informação nesse processo

São Paulo, 2020

RAFAEL JOAO DA SILVA

Batendo um papo com a informação: O uso dos *chatbots* para a recuperação da informação e a contribuição da Ciência da Informação nesse processo

Versão Corrigida (versão original disponível na Biblioteca da ECA/USP)

Dissertação apresentada à Escola de Comunicações e Artes da Universidade de São Paulo, para obtenção do título de Mestre em Ciências – Profissional

Área de Concentração: Organização, Mediação e Circulação da Informação

Orientação: Profa. Dra. Vânia Mara Alves Lima

São Paulo, 2020

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação na Publicação
Serviço de Biblioteca e Documentação
Escola de Comunicações e Artes da Universidade de São Paulo
Dados inseridos pelo(a) autor(a)

Silva, Rafael
Batendo um papo com a informação: O uso dos chatbots para a recuperação da informação e a contribuição da Ciência da Informação nesse processo. / Rafael Silva; orientadora, Vania Mara Alves Lima. -- São Paulo, 2020.
125 p.: il.
Dissertação (Mestrado) - Programa de Pós-Graduação em Ciência da Informação - Escola de Comunicações e Artes / Universidade de São Paulo.
Bibliografia
Versão corrigida
1. Chatbots 2. Inteligência Artificial 3. Organização da Informação 4. Recuperação da Informação 5. Ciência da Informação I. Mara Alves Lima, Vania II. Título.
CDD 21.ed. - 020

Elaborado por Alessandra Vieira Canholi Maldonado - CRB-8/6194

RAFAEL JOÃO DA SILVA

Batendo um papo com a informação: O uso dos *chatbots* para a recuperação da informação e a contribuição da Ciência da Informação nesse processo

Dissertação apresentada à Escola de Comunicações e Artes da Universidade de São Paulo, para obtenção do título de Mestre em Ciências – Profissional

Área de Concentração: Organização, Mediação e Circulação da Informação.

Orientação: Profa. Dra. Vânia Mara Alves Lima

Aprovado em:___/___/_____

BANCA EXAMINADORA

_____

Profa. Dra. Vânia Mara Alves Lima – Orientadora (USP)

_____

Prof. Dr. Rogério Aparecido Sá Ramalho (PPGCI/UFScar)

_____

Prof. Dr. José Eduardo Santarem Segundo (PPGCI/Unesp)

_____

DEDICATÓRIA

Dedico este estudo a minha família, amigos, professores e a todos os estudantes e profissionais que tenham interesse no tema.

AGRADECIMENTOS

Agradeço em especial a minha família, amigos e a todas as pessoas que direta ou indiretamente fizeram parte da minha vida e acreditaram em mim. A direção e administração da Escola de Comunicações e Artes (ECA) da Universidade de São Paulo (USP) pelas oportunidades e por proporcionarem um ambiente acadêmico propício a aprendizagem. Muito obrigado a todos os meus professores(as) pela troca de conhecimento e inspiração, em particular a minha orientadora pela parceria, correções e resiliência.

EPÍGRAFE

"Tudo no universo e tudo do homem seria registrado a uma distância assim como foi produzido. Desta maneira, uma imagem em movimento do mundo seria criada, um verdadeiro espelho de sua memória. De longe, todos serão capazes de ler o texto projetado em uma tela individual, ampliado ou limitado, do objeto desejado. Desta forma, cada um de sua cadeira será capaz de contemplar a criação, no todo ou em certas partes".[1]

Paul Otlet
(WRIGHT, 2014, p. 8)

[1] Paul Otlet, em 1934, profetizando o que seria, o hoje, a internet.

RESUMO

A preocupação com a recuperação da informação vem de épocas remotas e passou a ser amplamente abordada em diversas áreas do conhecimento. Essa preocupação ganhou amplitude devido a crescente quantidade de informações originadas e disponibilizadas pelo uso de novas Tecnologias da Informação e Comunicação (TIC). As TICs também proporcionam novas ferramentas de recuperação da informação, sendo uma delas os *chatbots*. Por meio dos métodos de pesquisa bibliográfica e estudos de casos múltiplos, a presente pesquisa tem por objetivo evidenciar como os *chatbots* estão colaborando para a recuperação da informação no âmbito da Ciência da Informação (CI) e, em contrapartida, como essa mesma Ciência ajuda no processo de melhoria e desenvolvimento desses sistemas. Desta forma, explora-se o campo da Ciência da Informação e suas vertentes. Versa-se sobre o que são os *chatbots* e suas principais tecnologias e tendências. Analisa-se alguns casos de uso desses agentes sob o ponto de vista de suas colaborações em facilitar a recuperação da informação. Por fim, nota-se como algumas ferramentas da CI contribuem para o desenvolvimento e melhorias desses sistemas. Como considerações finais, destaca-se que a Ciência da Informação e seus subdomínios são essenciais para auxiliar a arquitetura de projetos de inteligência artificial, como os *chatbots*. Trata-se de uma pesquisa exploratória que não tem o intuito de ser exaustiva, mas sim abrir caminhos para que novas discursões entre a academia e o mercado profissional reverberem a partir dos assuntos levantados aqui.

Palavras-Chave: Chatbots. Recuperação da Informação (RI). Inteligência Artificial (IA). Ciência da Informação (CI). Organização da Informação (OI).

ABSTRACT

The concern with information retrieval comes from remote times and has come to be widely addressed in several areas of knowledge. This concern became more widespread due to the growing amount of information originating and made available by the use of new Information and Communication Technologies (ICT). ICTs also provide new information retrieval tools, one of which is chatbots. Through methods of bibliographic research and multiple case studies, this research aims to show how chatbots are collaborating for the information retrieval in the scope of Information Science and how that same Science can help in process of improvement and development of these systems. Thus, the field of Information Science and its aspects is explored. It is analyzed what chatbots are and their main technologies and trends. Some use cases of these agents are examined from the point of view of their collaboration in facilitating information retrieval. Finally, it is evidenced how some Information Science tools can contribute to the development and improvement of these systems. As final considerations, it should be noted that Information Science and its subdomains are essential to assist in the architecture of artificial intelligence projects such as chatbots. This is an exploratory research that is not intended to be exhaustive, but rather opens ways for new discourses between academic and the market based on the topics raised here.

Keywords: Chatbots. Information Retrieval (IR). Artificial Intelligence (AI). Library and Information Science (LIS). Knowledge Organization (KO).

LISTA DE ILUSTRAÇÕES

Ilustração 1: Processo de geração de conhecimento 28
Ilustração 2: Estrutura do processo de Recuperação da Informação. 30
Ilustração 3: Estrutura de um sistema especialista 32
Ilustração 4: Sistemas de Organização do Conhecimento 38
Ilustração 5: Esquematização do Teste de Turing 39
Ilustração 6: Esquematização do funcionamento de um *chatbot* 41
Ilustração 7: Linha do tempo dos chatbots 43
Ilustração 8: Exemplo de intenção em uma frase 55
Ilustração 9: Exemplo de entidade em uma frase. 55
Ilustração 10: Exemplos de integrações em um sistema de *chatbot* 58
Ilustração 11: Ilustração de diálogo simples com a linguagem AIML 60
Ilustração 12: Aplicação da metodologia de *Sense making* para os *chatbots* 93
Ilustração 13: Levantamento de entidades e intenções no log de um atendimento 99
Ilustração 14: Exemplo de tesauro para o termo Smartphone 103
Ilustração 15: Taxonomia para um *Chatbot* de pedido de pizza 106
Ilustração 16: Ontologia para um *Chatbot* de Receitas 110

LISTA DE FIGURAS

Figura 1: Resultados de documentos acadêmicos com os termos “Ciência da Informação” e “*chatbots*” 17
Figura 2: Resultados de documentos acadêmicos com os termos “*Information science*” e *“chatbots”* 18
Figura 3: Disciplinas da Inteligência Artificial 46
Figura 4: Exemplo fictício de *chatbot* de viagens baseado em regras 51
Figura 5: Exemplo fictício de *chatbot* de viagens com Inteligência Artificial 52
Figura 6: Exemplo de persona (avatar) em um *chatbot* sobre saúde 54
Figura 7: Exemplo de uma estrutura de diálogo. 56
Figura 8: *Chatbot* do Ponto Frio utilizando o Facebook Messenger como canal 57
Figura 9: Telas de exemplos do IBM Watson Assistant 62
Figura 10: Telas de exemplos do DialogFlow 63
Figura 11: Algumas telas do ChatFuel 64
Figura 12: Gartner Hype Cycle para Inteligência Artificial 67
Figura 13: Momento que mostra a interação com o IBM Watson na Pinacoteca 71
Figura 14: Funcionamento do *chatbot* – Manual Cognitivo 73
Figura 15: Funcionamento da BIA no WhatsApp 75
Figura 16: Funcionamento do *chatbot* Gal no site da Gol 77
Figura 17: *Chatbot* Lu em funcionando no site da Magazine Luiza 79
Figura 18: *Chatbot* Judi em funcionamento no site do TJSP 81
Figura 19: Funcionando do CheckCorona no WhatsApp 82
Figura 20: Interação com a Bia via Facebook Messenger 84
Figura 21: *Chatbot* não reconhece a intenção de pergunta 86

SUMÁRIO

1. INTRODUÇÃO 13

1.1. Problema de pesquisa e justificativa 16

1.2. Objetivos 20

1.2.1. Objetivo geral 20

1.2.2. Objetivos específicos 20

1.3. Procedimentos metodológicos 21

2. A CIÊNCIA DA INFORMAÇÃO E A RECUPERAÇÃO DA INFORMAÇÃO 23

2.1. O que é Ciência da Informação? 24

2.2. Dado, informação e conhecimento 25

2.3. Recuperação da Informação 28

2.4. Sistemas de Organização do Conhecimento 34

3. BEM-VINDO À ERA DOS CHATBOTS 39

3.1. Sobre os *chatbots*: definição, história e evolução 40

3.2. O uso de Inteligência Artificial nos *chatbots* 44

3.3. Tipos, elementos e arquitetura dos *chatbots* 51

3.4. Desenvolvimento dos *chatbots* 59

3.5. Popularidade e principais tendências 65

4. BATENDO UM PAPO COM A INFORMAÇÃO: ESTUDOS DE CASOS 69

4.1. Quem é o autor da obra? – Estudo de Caso: A voz da Arte 69

4.2. O que significa esse sinal? – Estudo de Caso: Manual Cognitivo 72

4.3. Qual meu saldo atual? – Estudo de Caso: BIA do Bradesco 74

4.4. Posso levar isso na bagagem? – Estudo de Caso: Gal da Gol 76

4.5. Aceita cartão? – Estudo de caso: Lu da Magazine Luiza 78

4.6. Quero abrir um processo – Estudo de Caso: Judi do TJSP 80

4.7. Estou com o Coronavírus? – Estudo de Caso: CheckCorona 81

4.8. Você tem esse livro? – Estudo de Caso: Bia da Biblioteca da PUC 83

***5.* RELAÇÃO E CONTRIBUIÇÃO DA CI PARA OS *CHATBOTS* 86**

5.1. Estudos de usuário da informação para definição de persona 89

5.2. A indexação para identificar entidades e intenções 94

5.3. O uso dos tesauros na padronização de intenções e entidades 100

5.4. Organizando fluxos de diálogo com a Taxonomia 104

5.5. Utilizando a Ontologia no relacionamento de conversas 107

6. CONSIDERAÇÕES FINAIS 112

REFERÊNCIAS 115

1. INTRODUÇÃO

A sociedade vivenciou, ao longo de muitos anos, importantes avanços que proporcionaram cada vez mais acesso à informação, conhecimento e cultura; sobretudo pela facilidade proporcionada pelos avanços tecnológicos. Vive-se hoje a chamada Sociedade da Informação e do Conhecimento, onde informações são criadas e disseminadas a cada instante em diversos suportes (BARRETO, 2007).

Contudo, percebe-se que ao passo que mais informações são criadas, maior é o desafio de organizar e recuperar, tendo como objetivo de transformá-las em conhecimento ao usuário final. Nesse sentido, nos últimos anos, um novo fenômeno digital, chamado de *chatbots*, vem sendo adotado por diversas empresas e instituições por facilitar o acesso a informação por meio de uma plataforma intuitiva que utiliza o processamento de linguagem natural.

Os *chatbots* são sistemas/robôs/agentes inteligentes que têm como objetivo simular uma conversa natural com os usuários em torno de um ou vários domínios específicos, levando o ser humano, na maioria das vezes, a pensar que está conversando com uma outra pessoa. Essa interação é feita por meio de linguagem natural, utilizando-se de texto ou voz (CRUZ; ALENCAR; SCHMITZ, 2018). Deste modo, por meio desses sistemas, empresas estão permitindo que, literalmente, seus clientes tenham uma conversa natural com seus acervos de informações (banco de dados, obras, promoções etc.). No contexto atual, muitos desses *chatbots* são integrados a canais populares de conversação já existentes, como o WhatsApp, Messenger do Facebook, e-mails, SMS, entre outros.

Diversos *chatbots* estão sendo criados tendo como foco assuntos específicos. Um exemplo é o *chatbot* brasileiro chamado Henry (de Henry Ford). Por meio desse *chatbot*, o usuário interage, através de linguagem natural, com um banco de dados que contém informações sobre o universo automobilístico. O usuário consegue fazer perguntas sobre carros e o sistema é capaz de respondê-las de maneira lógica mediante uma conversa natural (PORTAL IDGNOW, 2017).

Quando o assunto é informação sobre saúde, existem diferentes *chatbots*. Um deles é o *Tá-na-Hora*. Por intermédio de SMS, as pessoas conseguem conversar e esclarecer dúvidas com um robô que fornece informações sobre prevenção de doenças, cuidados com a saúde e dicas de nutrição. O sistema ganhou o Prêmio Internacional do Global Steering Group por ajudar a

população da cidade de Marechal Deodoro a ter acesso a informações sobre temas como: prevenção de epidêmicas como leptospirose, cólera, febre, entre outras enfermidades comuns na região em épocas de enchentes (STARTSE, 2017).

Outro exemplo é o projeto chamado "A Voz da Arte", criado em 2017 pela IBM em parceria com a Pinacoteca de São Paulo (PORTAL ADNEWS, 2017). Utilizando o IBM Watson (Inteligência Artificial da IBM), criou-se um *chatbot*, de nome Isabela, que possibilitou que os visitantes interagissem, por meio de comando de voz e em linguagem natural, com algumas obras disponíveis na exposição, como Mestiço, de Cândido Portinari e São Paulo, de Tarsila do Amaral. Utilizando um *smartphone* e fones de ouvido – que ficavam próximos a essas obras -, os visitantes podiam fazer perguntas ao sistema como: "*Quem é o pintor?*" ou "*O personagem está com raiva?*" ou "*Qual escola esta obra pertence*?". As perguntas eram realizadas de diversas formas, sendo que o sistema era capaz de interpretá-las e proporcionar respostas em tempo real mediante banco de dados que foi alimentando durante 6 meses por curadores da Pinacoteca.

Não obstante, a área financeira também se beneficia com a utilização de *chatbots*. É o caso do *chatbot* do Banco Bradesco – a BIA (COMPUTERWORLD, 2019). Utilizando a plataforma WhatsApp ou o próprio aplicativo da instituição, os clientes do banco conseguem fazer perguntas relacionadas a abertura de contas, saldo, cadastramento de senhas. Além disso, uma pergunta chama a atenção dos usuários. Basta perguntar "*BIA, como está a economia hoje*" para que o agente consiga trazer respostas atualizadas sobre dados econômicos atuais, informações que podem ser fundamentais para as pessoas que têm dúvidas sobre onde investir o seu dinheiro. Em 2019, a plataforma já contava com mais de 100 milhões de conversas (COMPUTERWORLD, 2019), um número expressivo, considerando que o aplicativo foi oficialmente lançado para o público geral em 2017.

Observa-se nos exemplos que muitas das informações fornecidas por esses *chatbots* já constam nos sites das empresas proprietárias desses sistemas. Em outras palavras, com a utilização desses agentes, esses usuários não precisam navegar horas a fio em uma página para encontrar as informações desejadas. Portanto, bastaria perguntar para esses sistemas, utilizando linguagem natural, o que se deseja saber. Como o *chatbot* busca e entrega a informação desejada em tempo real por meio de interação natural, em uma primeira análise, pode-se inferir que eles facilitam a recuperação da informação para o usuário final.

Nesse sentido, os *chatbots* podem ser analisados sob o ponto de vista de diversas áreas do conhecimento, mas no âmbito da Ciência da Informação (CI) podem ser inseridos especificamente em discussões sobre a recuperação da informação. Por este motivo, o presente projeto de pesquisa se propõe a responder a seguinte pergunta: ***Como os chatbots estão colaborando para a recuperação da informação e qual a possível contribuição da Ciência da Informação no processo de melhoria e desenvolvimento desses agentes?***

Pretende-se, dessa maneira, explorar em quatro capítulos, a utilização de *chatbots* no sentido da Organização da Informação e do Conhecimento em duas vertentes: a primeira, a partir de estudos de casos, propõe analisar como esses sistemas contribuem para a recuperação da informação, área de estudo da Ciência da Informação; a segunda, tem por objetivo analisar a contribuição da Ciência da Informação no processo de desenvolvimento e melhoria desses sistemas com a utilização de ferramentas: estudo do usuário, indexação, controle de vocabulário, tesauros, ontologias.

O primeiro capítulo apresenta uma visão geral sobre a área da Ciência da Informação, sua história e seus principais conceitos. Explora-se sobre a diferença de Sociedade da Informação e Sociedade do conhecimento, bem como a distinção entre dado, informação e conhecimento. Por fim, define-se o que é Recuperação da Informação e como os Sistemas de Organização do Conhecimento são utilizados.

No segundo capítulo, busca-se detalhar o que são os *chatbots*, sua história e os principais atores nesse processo. Ao analisar a evolução e os tipos de *chatbots*, apresenta-se, de forma sucinta, o campo da Inteligência Artificial e suas vertentes, tais como: *Machine Learning*, *Deep Learning*, Linguagem Natural. Explora-se ainda os elementos de um *chatbot*, definindo termos como: base de conhecimento, diálogos, intenção e entidades. Por fim, busca-se compreender as principais plataformas de desenvolvimentos desses agentes e como eles se conectam com outros sistemas.

Por meio da análise de 8 estudos de casos, o terceiro capítulo demonstra como esses agentes estão colaborando para a recuperação da informação e o acesso ao conhecimento através dos seguintes casos: o projeto da IBM na Pinacoteca de São Paulo para a interação com obras de arte; o manual cognitivo da Volkswagen, projeto de *chatbot* para facilitar a interação e

utilização dos carros da marca; A BIA do Bradesco, como caso de sucesso para obtenção de informações financeiras; a GAL, *chatbot* da Gol Linhas Aéreas utilizado para informações sobre as viagens aéreas; a Lu, nome dado ao *chatbot* da Magazine Luiza que auxilia os clientes com informações de como realizar compras no site da empresa; A Judi, do Tribunal de Justiça de São Paulo, que auxilia com informações de abertura de processo; o CheckCorona, projeto criado com o intuito de auxiliar os usuários sobre os sintomas da Covid-19 e, por fim, a BIA, *chatbot* criado pela PUC para o auxílio no serviço de referência da biblioteca.

Por fim, o último capítulo discute como a Ciência da Informação colabora para a melhoria e o desenvolvimento dos *chatbots*. O capítulo versa a importância de organizar o conhecimento para alimentar sistemas de inteligência artificial, como os *chatbots*, e, nesse ínterim, sugere-se algumas ferramentas da CI como forma de facilitar esse processo. Para isso, inicia-se analisando como os estudos de usuários da informação podem contribuir para a definição de audiência e personas para os *chatbots*. Em seguida, é exemplificado como a Indexação pode ajudar na identificação de entidades e intenções. Por último, as possíveis aplicações de alguns Sistemas de Organização do Conhecimento são investigadas no desenvolvimento de um *chatbot*, por exemplo, as taxonomias, tesauros e ontologias -, sobretudo no que tange a organização de fluxo de diálogos.

1.1. Problema de pesquisa e justificativa

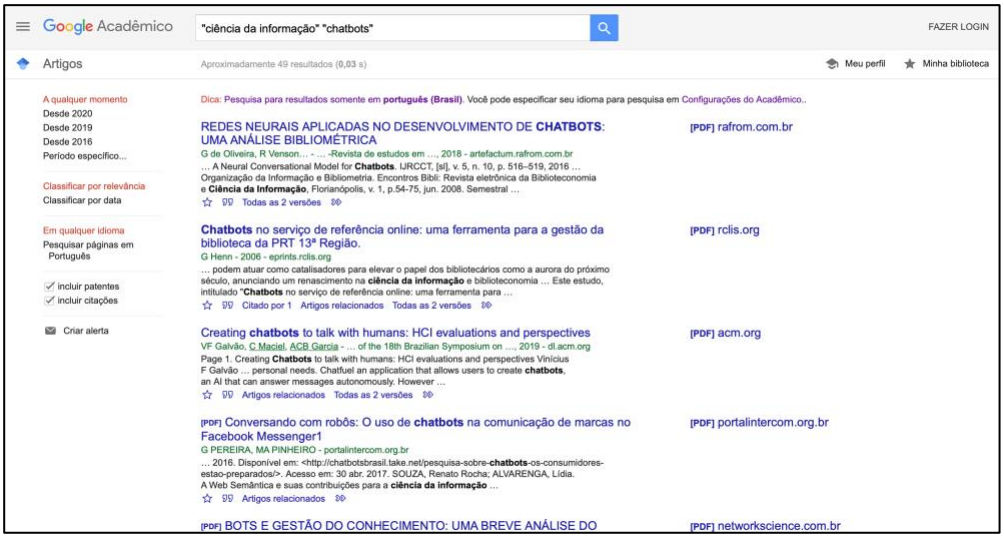
A possibilidade de ter uma máquina capaz de interagir com o ser humano tem sido, há muito tempo, alvo de diversas pesquisas na área de Inteligência Artificial e, nos últimos anos, isso tem se tornado cada vez mais possível e acessível (CRUZ; ALENCAR; SCHMITZ, 2018). Os *chatbots* são resultados dessas pesquisas. Hoje, é possível criar, de forma gratuita e em pouco tempo, sistemas capazes de interpretar um grande número de informações e disponibilizar conhecimento de forma acessível para os usuários, por meio de um diálogo em linguagem natural.

Os *chatbots* podem ser analisados sob o ponto de vista de diversas áreas do conhecimento, principalmente no que se refere à recuperação da informação no âmbito da Ciência da Informação. Porém, são poucas as pesquisas realizadas na área com o foco em entender esse fenômeno no contexto atual e suas possibilidades, tanto no que tange como esses sistemas

podem contribuir para a CI e, principalmente, o inverso, como a CI pode contribuir para a melhoria e desenvolvimento desses sistemas. A falta de pesquisa nesse contexto é compreensível, já que os *chatbots* começaram a popularizar-se apenas nos últimos anos e, além disso, é um tópico de maior interesse para a área da Ciência da Computação.

Essa carência de pesquisas na área da Ciência da Informação envolvendo *chatbots* fica evidente quando se realiza uma pequena busca com os termos em português "Ciência da Informação" e "*chatbots*" utilizando o Google Acadêmico e princípios da revisão sistemática (MANCINI, 2007) . Observa-se um retorno de 49 resultados (envolvendo trabalhos de graduação, artigos, teses), conforme a Figura 1.

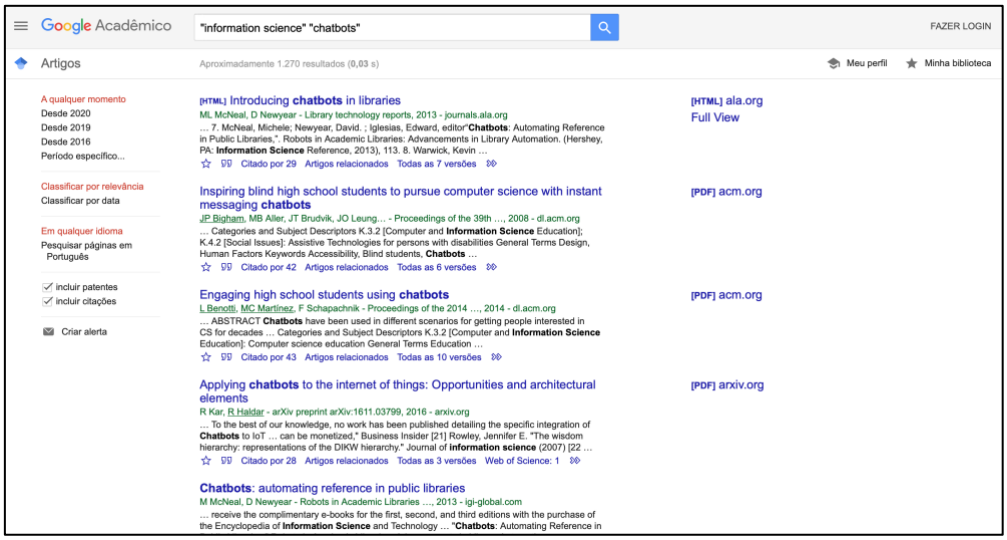
Figura 1: Resultados de documentos acadêmicos com os termos "Ciência da Informação" e "*chatbots*"



Fonte: Google Acadêmico – Pesquisa realizada em 10/02/2020

Já ao pesquisar pelos mesmos termos, porém em inglês, o resultado é bem mais expressivo, algo em terno de 1.270 resultados (vide Figura 2). Entretanto, ao analisar o teor desses resultados, nota-se ainda uma carência em explorar as contribuições da CI para os *chatbots* e vice-versa.

Figura 2: Resultados de documentos acadêmicos com os termos "*Information science*" e "*chatbots*"



Fonte: Google Acadêmico – Pesquisa realizada em 10/02/2020

Com base nos estudos de revisão sistemática (MANCINI, 2007), as palavras chaves utilizadas nessa busca foram escolhidas por representar a pergunta problema desse projeto. Obviamente, existem expressões que possuem sinônimos próximos à essas palavras, como Processamento de Linguagem Natural (como equivalência a chatbot). Entretanto, esse é apenas um dos aspectos de um chatbot e, portanto, não é um termo que abarca por completo os assuntos discutidos aqui. Além disso, foi utilizado o operador booleano de aproximação – as "aspas" - para limitar a busca e garantir que os resultados retornassem documentos que tivessem relação com ambos os termos pesquisados, e não um dos termos isolados – o que, evidentemente, traria um número maior de documentos encontrados. O Google Acadêmico foi utilizado como repositório de busca pela sua facilidade de ser um serviço que indexa diversas bases de dados científicas em um único lugar (MUGNAINI; STREHL, 2008)

Com isso, apesar do pouco material sobre o assunto, ainda é possível analisar algumas pesquisas na área da Ciência da Informação relacionadas aos *chatbots*. Há estudos que procuram explorar o uso de *chatbots* como serviços de referência em unidades de informação, como o artigo de Ataíde et al. (2010) "***Tecnologia Da Informação E Serviços De Referência Eletrônicos: Uma Proposta de Aplicação Baseada em Chatbots e Ontologias***" que analisa a utilização de *chatbots* como serviço de referência de uma maneira mais abrangente; ou o artigo de Oliveira (2010) "***Dr. Pierre: Um Chatbot com Intenção e Personalidade Baseado em Ontologias para***

Apoiar o Ensino de Psiquiatria" que explora a criação de um *chatbot* com base no uso de ontologias.

Entretanto, percebe-se escassez no sentido de explorar casos reais e de como esses sistemas estão sendo utilizados no contexto atual para a recuperação da informação, sobretudo quando esses sistemas utilizam tecnologias oriundas da Inteligência Artificial. Além disso, demanda explorar como a Ciência da Informação pode contribuir para a melhoria e desenvolvimento desses sistemas, já que, ao interagir com alguns desses agentes, percebe-se a falta de um controle de vocabulário e a relação entre os termos. Nesse contexto, a CI demonstra ser uma ciência essencial para mitigar esses problemas.

Explorar a contribuição desses sistemas para a área da Ciência da Informação é algo que parece ser justificável, já que os *chatbots* são desenvolvidos na Ciência da Computação, área que atualiza e compartilha processos fundamentais para a Ciência da Informação, em especial aqueles que tangem pesquisas relacionadas as técnicas que automatizam o gerenciamento e recuperação da informação (CAFEZEIRO; COSTA; KUBRUSLY, 2016). Entretanto, o outro lado também é tão importante quanto. Isto é, entender como a Ciência da Informação pode colaborar para o uso e desenvolvimento desses sistemas é algo essencial, particularmente quando se observa que a fundamentação de muitas tecnologias da informação tem contribuições importantes que se originaram primeiramente na Ciência da Informação (CAFEZEIRO; COSTA; KUBRUSLY, 2016). Isso fica evidente quando se analisa as seguintes citações:

> Embora a Inteligência Artificial esteja recebendo muita visibilidade, o fato de essa tecnologia exigir algum elemento da engenharia do conhecimento, arquitetura da informação e fontes de dados de alta qualidade não é bem conhecido. [...] a maioria desses sistemas requerem uma quantidade significativa de trabalho por parte dos seres humanos antes de redes neurais, aprendizado de máquina e processadores de linguagem natural possam trabalhar sua magia (EARLEY, 2016).

> Analistas na Gartner preveem uma taxa de falha de 99% iniciativas de inteligência artificial no gerenciamento de serviços de TI, devido à falta de práticas estabelecidas de gestão do conhecimento. De fato, sem acesso a um bom conteúdo, os *chatbots* tem muito pouco para trabalhar durante as interações com os usuários (MINDTOUCH, 2019).

> Os seres humanos criam conhecimento, enquanto máquinas processam, armazenam e agem nesse conhecimento. A Inteligência Artificial (IA) é aplicado ao conhecimento humano. Para isso, as organizações precisam desenvolver as fundações para avançar a IA, isto inclui a captura e curadoria desse conhecimento e a construção de informações estruturadas que formam o andaime para esse conhecimento. Sem esses componentes, os algoritmos não têm nada o que fazer. (EARLEY, 2016).

Com essas citações, observa-se que, embora a promessa de que sistemas com inteligência artificial (como os *chatbots*) sejam tendências, sobretudo para a contribuição do acesso a informação, nada disso funciona "magicamente" sem uma organização adequada das informações processadas por esses algoritmos. Assim dizendo, toda a estrutura - do que se chama de inteligência – utilizada por esses sistemas, ainda é desenvolvida de forma manual pelo ser humano. Dessa forma, para que esses sistemas funcionem de maneira correta e que consigam satisfazer as necessidades de informação dos usuários, é necessário que a sua fundação seja desenvolvida baseada em: organização da informação, entendimento da necessidade de busca do usuário, análise e curadoria dos conteúdos apresentados por esses sistemas, assim por diante. Nesse sentido, nota-se que a Ciência da Informação tem um papel muito importante na base e desenvolvimento desses agentes inteligentes.

1.2. Objetivos

Esta pesquisa propõe-se a alcançar os seguintes objetivos:

1.2.1. Objetivo geral

Analisar a utilização dos *chatbots* para a recuperação da informação e identificar as contribuições da CI no processo de melhoria e desenvolvimento desses agentes.

1.2.2. Objetivos específicos

- Analisar os *chatbots* (origem, tendências, desafios, arquitetura e uso de Inteligência Artificial);
- Identificar os processos e instrumentos para organização, representação e recuperação na Ciência da Informação;
- Levantar e analisar, no contexto atual, casos reais de utilização de *chatbots* para a recuperação da informação;
- Analisar como, e se, os processos e instrumentos para a organização, representação e recuperação na Ciência da Informação podem contribuir para o desenvolvimento e melhorias dos *chatbots*.

1.3. Procedimentos metodológicos

Com base nos objetivos de estudo e problema apresentados para este trabalho, escolheu-se a pesquisa exploratória como o tipo de pesquisa mais adequada para o assunto tratado aqui. Segundo Gil (2002, pg. 41), pesquisas exploratórias têm por objetivo:

> [...] proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses. Pode-se dizer que estas pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições. Seu planejamento é, portanto, bastante flexível, de modo que possibilite a consideração dos mais variados aspectos relativos ao fato estudado.

Já Mueller (2016, pg. 25) diz que o objetivo da pesquisa exploratória é:

> [...] reunir dados, informações, padrões, ideias ou hipóteses sobre um problema ou questão de pesquisa com pouco ou nenhum estudo anterior. Esse tipo de pesquisa não tem o objetivo de tentar testar uma hipótese, mas de procurar padrão.

Portanto, é notório que a pesquisa exploratória é o modelo de pesquisa que mais se adequa ao presente trabalho, pois o assunto tratado aqui é relativamente novo e pouco explorado, sobretudo na área da Ciência da Informação. Com isso, como o próprio objetivo geral da pesquisa evidência, trata-se de um trabalho que busca explorar e entender possibilidades dentro da temática dos *chatbots*, tendo como finalidade levantar informações que sirvam para que novos estudos possam ser realizados com base nos resultados aqui apresentados.

Além disso, no que tange a sua abordagem, esta é uma pesquisa qualitativa, pois, de acordo com Muller (2016, pg. 27) trata-se de uma abordagem aplicada às áreas com pouco conhecimento teórico ou conceitual – como é o caso dos *chatbots*. É um tipo de abordagem que ajuda a compreender o objeto de estudo e a construí-lo a partir de novos aspectos e sob novas perspectivas. É, portanto, uma abordagem apropriada quando analisado os objetivos desta pesquisa, bem como seu objeto de estudo e problema.

No que se refere aos métodos utilizados para a coleta e análise de dados, foram escolhidas as seguintes técnicas: pesquisa bibliográfica e estudos de casos múltiplos.

De acordo com Severino (2017, pg. 131), a pesquisa bibliográfica é:

> [...] aquela que se realiza a partir do registro disponível, decorrente de pesquisas anteriores, em documentos impressos, como livros, artigos, teses etc. Utiliza-se de

> dados ou de categorias teóricas já trabalhados por outros pesquisadores e devidamente registrados. Os textos tornam-se fontes dos temas a serem pesquisados. O pesquisador trabalha a partir das contribuições dos autores dos estudos analíticos constantes dos textos.

Para esta pesquisa, a revisão bibliográfica será realizada de maneira a ampliar e aprofundar a análise dos principais conceitos levantados até o momento, como: ciência da informação, sistemas de organização do conhecimento, usuário da informação, análise de conteúdo, terminologia, indexação, representação e recuperação da informação, *chatbot*, Inteligência Artificial (*Machine Learning*, *Deep Learning* e Processamento de Linguagem Natural).

Em relação ao Estudo de Caso, Gil (2002, pg. 52) o define como:

> O estudo de caso é uma modalidade de pesquisa amplamente utilizada nas ciências biomédicas e sociais. Consiste no estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita seu amplo e detalhado conhecimento, tarefa praticamente impossível mediante outros delineamentos já considerados.

Desse modo, tem-se como objetivo utilizar os estudos de caso múltiplos como técnica para contextualizar o estudo teórico e analisar a contribuição dos *chatbots* para a Ciência da Informação. Mueller (2007, pg.46) explica que os estudos de casos múltiplos têm como finalidade analisar mais de um caso, sendo que as provas resultantes desses estudos são consideradas mais convincentes, pois normalmente são análises mais robustas e rigorosas que utilizam de uma abordagem mais completa. Com isso, por meio dessa técnica de pesquisa, almeja-se analisar e descrever casos atuais de usos de *chatbots* no âmbito da CI, sob o prisma de entender como facilitam a recuperação da informação.

Em suma, é um estudo exploratório com linguagem acessível e que não pretende ser exaustivo, mas sim levantar possibilidades para que novos estudos e diálogos surjam entre a academia e o mercado, com base nos tópicos e evidências que serão abordados neste trabalho.

2. A CIÊNCIA DA INFORMAÇÃO E A RECUPERAÇÃO DA INFORMAÇÃO

Desde o surgimento da escrita, o homem sempre procurou maneiras de registrar e disseminar o conhecimento. Burke (2003, 2012) em seus livros - *Uma história Social do Conhecimento I e II* - descreve muito bem essa trajetória. Segundo o autor, desde muito cedo, a sociedade busca criar mecanismos para o registro de informações e, além disso, formas de tornar o conhecimento acessível para todos. Isso fica evidente com a criação da prensa de Gutenberg, o surgimento das instituições e das primeiras bibliotecas, o advento das enciclopédias e, por fim, a chegada da internet (BARRETO, 2007). Hoje, com a facilidade ao acesso a informação, é muito comum dizer que se vive na Sociedade da Informação e do Conhecimento (BURCH, 2005).

Werthein (2000) descreve a Sociedade da Informação como:

> A expressão "sociedade da informação" passou a ser utilizada, nos últimos anos desse século, como substituto para o conceito complexo de "sociedade pós-industrial" e como forma de transmitir o conteúdo específico do "novo paradigma técnico-econômico". A realidade que os conceitos das ciências sociais procuram expressar refere-se às transformações técnicas, organizacionais e administrativas que têm como "fator-chave" não mais os insumos baratos de energia – como na sociedade industrial – mas os insumos baratos de informação propiciados pelos avanços tecnológicos na microeletrônica e telecomunicações.

Na Sociedade da Informação, as informações são geradas e disponibilizadas a todo o instante em uma quantidade cada vez maior e crescente. As pessoas possuem maior facilidade em consumir diversos tipos de conteúdo informacional, seja via internet, livros, bibliotecas e outros. Esse “boom informacional” foi impulsionado especialmente pelas novas Tecnologias de Informação e Comunicação (TICs), que proporcionaram um acesso mais dinâmico e rápido para o consumo da informação (BURCH, 2005). Entretanto, nesse contexto, cabe uma reflexão: será que hoje as pessoas também vivem em uma Sociedade do Conhecimento? Barreto (2007), indaga que:

> O ideal compartilhado seria o de se construir uma sociedade do conhecimento não só uma sociedade da informação. É um erro confundir a sociedade da informação com a sociedade do conhecimento. A sociedade da informação é uma utopia de realização tecnológica e a do conhecimento uma esperança da realização do saber.

Portanto, embora haja muita informação na atualidade (Sociedade da Informação), não se pode dizer que a sociedade está de fato se apropriando desta informação (Sociedade do

Conhecimento). Isto é, será que as pessoas realmente possuem um conhecimento prévio para entender as informações que consomem? As pessoas realmente estão encontrando o que buscam? O acesso à essas informações está sendo facilitado? Essa informação está organizada? Essas indagações são pertinentes, em especial quando se observa a importância da informação para o desenvolvimento do homem em sociedade, como observa Barreto (2007):

> A informação sintoniza o mundo, pois referencia o homem ao seu passado histórico, às suas cognições prévias e ao seu espaço de convivência, colocando-o em um ponto do presente, com uma memória do passado e uma perspectiva do futuro; o indivíduo do conhecimento se localiza no presente contínuo que é o espaço de apropriação da informação

Um campo de estudos que se preocupa com estas questões parece ser algo fundamental, já que, como mencionado na citação, a informação é um ativo importante para a evolução das pessoas em sociedade (BARRETO, 2007; SMIT 2012). É por meio da informação que o homem consegue desenvolver-se como indivíduo e, por conseguinte, transformar a sociedade. É nesse contexto, que a Ciência da Informação se configura em uma importante área de conhecimento ao ter a informação como o seu objeto de estudo.

2.1. O que é Ciência da Informação?

A CI surgiu como ciência na década de 50 (BARRETO, 2007; FREIRE, 2006). Entretanto, Barreto (2007) lembra que o livre fluxo de informação e sua distribuição tem sido um sonho da sociedade que transpassa diferentes épocas. Segundo ele, a luta por uma distribuição adequada do conhecimento produzido pela sociedade vem desde o século XVIII, passando por distintas gerações e grupos. Em consequência, múltiplos mecanismos foram criados para que esse conhecimento se tornasse mais acessível às pessoas; vide por exemplo, a construção da enciclopédia de Diderot e D'Alembert. A seguir algumas definições do que significa Ciência da Informação.

Para Borko (1968):

> Ciência da Informação é a disciplina que investiga as propriedades e o comportamento informacional, as forças que governam os fluxos de informação, e os significados do processamento da informação, visando à acessibilidade e a usabilidade ótima. A Ciência da Informação está preocupada com o corpo de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação, e utilização da informação. Isto inclui a pesquisa sobre a

> representação da informação em ambos os sistemas, tanto naturais quanto artificiais, o uso de códigos para a transmissão.

Saracevic (1996), define a CI como:

> A Ciência da Informação é um campo dedicado às questões científicas e à prática profissional voltadas para os problemas da efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação. No tratamento destas questões são consideradas de particular interesse as vantagens das modernas tecnologias informacionais.

Já Silva (2006) sugere uma definição para a CI com a ideia de info-comunicação, onde mescla as palavras informação e comunicação:

> Ciência da Informação é uma ciência social que investiga os problemas, temas e casos relacionados com o fenômeno info-comunicacional perceptível e cognoscível através da confirmação ou não das propriedades inerentes à gênese do fluxo, organização e comportamentos informacionais (origem, coleta, organização, armazenamento, recuperação, interpretação, transformação e utilização da informação).

Por fim, Capurro e Hjorland (2007) definem a CI da seguinte forma:

> Se ocupa com a geração, coleta, organização, interpretação, armazenamento, recuperação, disseminação, transformação e uso da informação, com ênfase particular, na aplicação de tecnologias modernas nestas áreas. Como uma disciplina, procura criar e estruturar um corpo de conhecimentos científico.

Com essas definições, percebe-se que desde a década de 1960, a Ciência da Informação se preocupa intrinsecamente em aprimorar e facilitar o acesso a informação para as pessoas. Como lembra SMIT (2012): "o principal objetivo da Ciência da Informação, a saber, contribuir para a melhoria das condições de vida do homem e da sociedade através do acesso à informação". A CI se configura como mediadora entre o usuário e a informação, preocupando-se com a usabilidade de sistemas, desenvolvendo técnicas para facilitar a recuperação da informação e, sobretudo, procurando caminhos para que a informação seja um importante ativo para a evolução do homem na sociedade por meio do conhecimento.

2.2. Dado, informação e conhecimento

Conforme Capurro e Hjorland (2007), por sua diversidade, informação é um termo bastante complexo, pois é utilizado em diversos contextos e em diferentes áreas, adotando significados distintos. Para a Ciência da Informação (CI), entender a definição de informação é imprescindível visto que se trata de seu objeto de pesquisa. Assim, compilou-se alguns autores

que têm se ocupado em definir o que é a informação nas últimas décadas.

Rafael Capurro e Hjorland (2007) analisaram o conceito informação desde suas origens latinas e gregas até a época atual. É interessante observar que o significado de informação passou por diversas mudanças no decorrer do tempo, influenciado sempre pelo pensamento da sociedade da época. Capurro e Hjorland (2007) versam que o conceito informação é interdisciplinar, ou seja, cada área do conhecimento tem uma interpretação diferente sobre o seu conceito e a problematiza com diferentes abordagens. Diante disso, fica a pergunta: qual seria a definição correta de informação para a CI?

Para responder à pergunta, Smit (2012) define o significado de informação dentro da CI enquanto um registro institucionalizado, isto é, para a autora a informação deve ser registrada para que ela sobreviva no tempo e no espaço. Portanto, se faz necessário o seu registro (seja ele físico ou eletrônico) e, ao mesmo tempo, essa informação precisa ser institucionalizada para que tenha um "selo de qualidade" – termo usado pela autora para designar a credibilidade da informação (SMIT, 2012). Em outras palavras, a informação precisa ser registrada e institucionalizada para carregar a potencialidade de utilidade informacional.

Já Buckland (1991) lembra que, a palavra informação é entendida como: estar informado; redução da ignorância ou incerteza. Entretanto, tal definição é irônica, já que o próprio termo é algo ambíguo e usado de diferentes maneiras. O autor define informação enquanto processo, conhecimento e coisa (objeto):

> Informação-como-processo: Quando alguém é informado, aquilo que conhece é modificado. Nesse sentido "informação" é "o ato de informar ...; comunicação do conhecimento ou "novidade" de algum fato ou ocorrência; a ação de falar ou o fato de ter falado sobre alguma coisa".
>
> Informação-como-conhecimento: "Informação" é também usado para denotar aquilo que é percebido na "informação-como-processo": o "conhecimento comunicado referente a algum fato particular, assunto ou evento; aquilo que é transmitido, inteligência, notícias". A noção de que informação é aquela que reduz a incerteza poderia ser entendida como um caso especial de "informação-como-conhecimento". Às vezes aumenta a informação incerteza
>
> Informação-como-coisa: O termo "informação" é também atribuído para objetos, assim como dados para documentos, que são considerados como "informação", porque são relacionados como sendo informativos, tendo a qualidade de conhecimento comunicado ou comunicação, informação, algo informativo.

Como visto, o termo informação emprega diversos significados, a depender do seu contexto de uso. Entretanto, para uma ciência que tem como objeto de pesquisa o fenômeno informação, se faz

necessário entendê-lo para que os resultados desses estudos estejam alinhados a ele. Nesse sentido, para o objetivo desta pesquisa, a definição de informação utilizada será a de informação-como-coisa, pois como propõe SMIT (2012), ela pode ser registrada e institucionalizada. Nesse ínterim, convém definir o que é dado e conhecimento, pois são dois termos que, muitas vezes, são utilizados como significados de informação, mas que para essa pesquisa, se faz necessário diferenciá-los. Comecemos pela definição de dados. Segundo Setzer (1998), dados pode ser definido como:

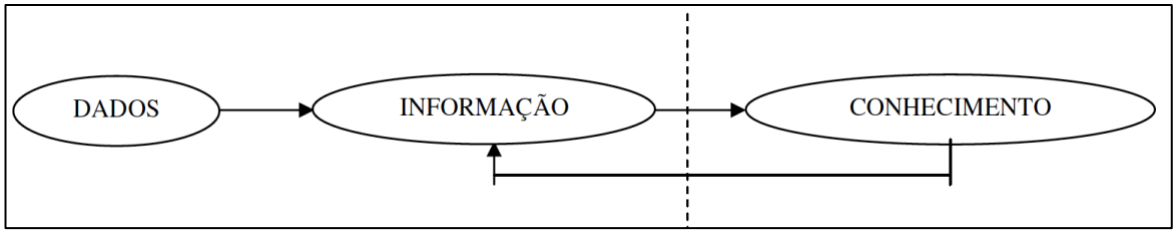
> [...]uma sequência de símbolos quantificados ou quantificáveis. Portanto, um texto é um dado. De fato, as letras são símbolos quantificados, já que o alfabeto, sendo um conjunto finito, pode por si só constituir uma base numérica (a base hexadecimal emprega tradicionalmente, além dos 10 dígitos decimais, as letras de A a E). Também são dados fotos, figuras, sons gravados e animação, pois todos podem ser quantificados a ponto de se ter eventualmente dificuldade de distinguir a sua reprodução, a partir da representação quantificada, com o original. É muito importante notar-se que, mesmo se incompreensível para o leitor, qualquer texto constitui um dado ou uma sequência de dados. Isso ficará mais claro no próximo item. (SETZER, 1998)

Já conhecimento é definido como:

> [...]o conhecimento não pode ser descrito; o que se descreve é a informação. Também não depende apenas de uma interpretação pessoal, como a informação, pois requer uma vivência do objeto do conhecimento. Assim, o conhecimento está no âmbito puramente subjetivo do homem ou do animal. Parte da diferença entre estes reside no fato de um ser humano poder estar consciente de seu próprio conhecimento, sendo capaz de descrevê-lo parcial e conceitualmente em termos de informação, por exemplo, através da frase "eu visitei Paris, logo eu a conheço" (supondo que o leitor ou o ouvinte compreendam essa frase). (SETZER, 1998)

Em suma, dados são símbolos e representações sem contexto que necessitam ser processados, tratados e interpretados para tornarem informação. Logo, informação são estruturas com significados que têm a capacidade de gerar conhecimento para uma pessoa. Conhecimento é algo que altera o cognitivo de um sujeito em contato com a informação. A qualidade desse novo conhecimento depende do contexto atual do sujeito, bem como suas vivências e experiências. Esse conhecimento adquirido pode se transformar em uma nova informação, agora remodelada e baseada nas experiências desse indivíduo. Um ciclo entre informação e conhecimento é gerado (vide Ilustração 1).

Ilustração 1: Processo de geração de conhecimento



Fonte: De Pontes Junior; De Aquino Carvalho; Azevedo (2013)

Como observado, o termo informação é muito mais complexo do que se entende casualmente e precisa ser visto como algo além de informar, como lembram Capurro e Hjorland (2007): "A coisa mais importante em CI é considerar a informação como uma força constitutiva na sociedade e, assim, reconhecer a natureza teleológica dos sistemas e serviços de informação". Desta forma, a CI se torna uma área do saber indispensável para propor soluções e reflexões que, indubitavelmente, contribuirão para que a sociedade se torne cada vez mais uma sociedade do conhecimento e que sistemas, como *chatbots*, sejam mais eficientes no processo de recuperação da informação.

2.3. Recuperação da Informação

A Recuperação da Informação (RI) é uma área estudada tanto na Ciência da Computação, como na Ciência da Informação. Para CI, a Recuperação da Informação é um dos seus pilares fundamentais, pois é nessa área que estão concentrados os estudos referentes aos processos e técnicas que têm por objetivo atender às necessidades de informação do usuário. Com um viés mais voltado à área da Ciência da Computação, os autores Baeza-Yates e Ribeiro-Neto (2013) no livro "Recuperação de Informação – Conceitos e Tecnologias das Máquinas de Busca", definem Recuperação da Informação da seguinte forma:

> A Recuperação de Informação (RI) é uma área abrangente da Ciência da Computação que se concentra principalmente em prover aos usuários acesso fácil as informações de seu interesse. [...] A recuperação de Informação trata da representação, armazenamento, organização e acesso a itens de informação, como documentos, páginas Web, catálogos online, registros estruturados e semiestruturados, objetivos multimídias, etc. A representação e a organização dos itens de informação devem fornecer aos usuários facilidade de acesso as informações de seu interesse.

Já com uma abordagem mais direcionada a Ciência da Informação, FERNEDA (2003) define a RI como:

> No contexto da Ciência da Informação, o termo "recuperação de informação" significa, para uns, a operação pela qual se seleciona documentos, a partir do acervo, em função da demanda do usuário. Para outros, "recuperação de informação" consiste no fornecimento, a partir de uma demanda definida pelo usuário, dos elementos de informação documentária correspondentes. O termo pode ainda ser empregado para designar a operação que fornece uma resposta mais ou menos elaborada a uma demanda, e esta resposta é convertida num produto cujo formato é acordado com o usuário (bibliografia, nota de síntese, etc.). Há ainda autores que conceituam a recuperação de informação de forma muito mais ampla, ao subordinar à mesma o tratamento da informação (catalogação, indexação, classificação).

Apesar do conceito ser utilizado a bastante tempo, o termo "*Information Retrieval*" (Recuperação da Informação) foi criado em 1951 por Calvin Mooers (FERNEDA, 2003). Segundo Mooers (1951 apud FERNEDA, 2003), a RI trata dos aspectos intelectuais da descrição da informação e sua especificação para busca, e também de qualquer sistema, técnicas ou máquinas que são empregadas na realização dessa operação.

A preocupação com a recuperação da informação vem de épocas remotas e passou a ser amplamente abordada em diversas áreas do conhecimento. Entretanto, essa preocupação ganhou mais amplitude devido a quantidade crescente de informações oriundas e disponibilizadas pelo uso de novas Tecnologias da Informação de Conhecimento – TICs (ALVES, Rachel C. V. et al, 2007). Os primeiros estudos em relação a recuperação da informação foram desenvolvidos de forma mais expressiva no final do século XIX, sob a ótica de diversos estudiosos. Os mais destacados, entretanto, são os estudos do cientista belga Paul Otlet e o cientista americano Vannevar Bush.

Na década de 1930, Paul Otlet, conhecido como o percursor da documentação, foi um visionário em sua época, pois mesmo com pouca tecnologia, ele foi capaz de sistematizar e elaborar instrumentos na tentativa de solucionar questões de recuperação e acesso ao conhecimento da época. Destaca-se a Classificação Decimal Universal (CDU), que é amplamente utilizada em diversas bibliotecas ao redor do mundo por facilitar a indexação e recuperação de informações. (TÁLAMO, 2007).

> [...] pode-se dizer que a grande contribuição de Otlet foi o enfoque dado ao conteúdo informacional e às técnicas e métodos de tratamento informacional desenvolvidas para a disponibilização, acesso e recuperação da informação. (ALVES, Rachel C. V. et al, 2007).

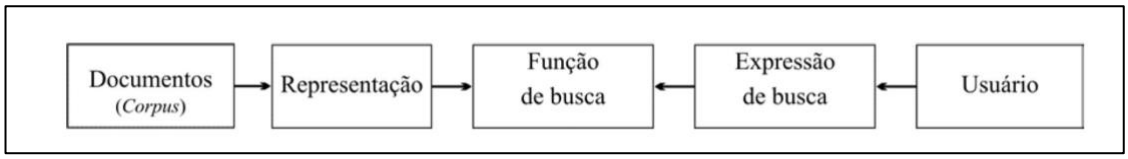
Igualmente Otlet, o americano Vannevar Bush, na década de 1940, preocupou-se em unir a ciência e a tecnologia com o intuito de tornar o conhecimento humano mais acessível. Ele foi responsável por diagnosticar a dificuldade de produção, organização, acesso, recuperação e

disseminação da informação impulsionado pelo crescimento do conhecimento humano na época. (FERNEDA, 2003; BARRETO, 2007).

> Já naquela época Bush abordou preocupações atuais da comunidade científica e da humanidade como um todo: Qual a maneira eficiente de armazenar e recuperar o conhecimento humano? Como fazer uso apropriado do crescente volume de informações? Como impedir que uma informação relevante passe despercebida em meio a uma abundante geração de conhecimento? Como mecanizar o registro de ideias e experiências de forma a não estacionar no tempo em função da limitada capacidade de memória? (ALVES, Rachel C. V. et al, 2007).

Bush idealizou o Memex (*Memory Extension*), um mecanismo que tinha como propósito automatizar as ações de organização e recuperação da informação. A ideia por trás desse mecanismo era reproduzir a capacidade de associação das ideias humanas. Com isso, seria possível armazenar livros, registros, fotos e anotações que poderiam ser acessadas com extrema facilidade e velocidade, simulando a extensão da memória humana. Infelizmente, o Memex não se tornou realidade, já que não existia, naquela época, tecnologia para suportar essa ideia. A visão de Bush foi de extrema importância para a área de Ciência da Informação, sobretudo no que tange a recuperação da informação. (ALVES, Rachel C. V. et al, 2007). O processo de um Sistema de Recuperação da Informação (SRI) pode ser resumido conforme a Ilustração 2.

Ilustração 2: Estrutura do processo de Recuperação da Informação.



Fonte: Ferneda (2003)

Primeiro tem-se o *Documento (Corpus)* original que será recuperado de acordo com a busca do usuário. Para que esse documento possa ser encontrado e processado pelo sistema, é necessário que ele passe por um processo de *representação*, que consiste em resumir o conteúdo do documento original e definir pontos (termos) de acesso para a busca. Por meio de uma interface, o usuário insere a sua intenção de busca (*expressão de busca*) para logo em seguida o sistema transformá-la em uma *função de busca*, sendo assim, o texto inserido pelo usuário é convertido em uma estrutura compreendida pelos algoritmos para que a informação possa ser encontrada e disponibilizada.

Vale ressaltar que um sistema de recuperação da informação, como um *chatbot*, tem por objetivo retornar itens que atendam de fato as necessidades de informação do usuário. Trata-se, portanto, de um processo no qual busca encontrar uma informação que seja relevante de tal forma que possa contribuir para mudar o cognitivo da pessoa (gerar um novo conhecimento), conforme Ferneda (2003) elucida:

> O usuário de um sistema de recuperação de informação está, portanto, interessado em recuperar "informação" sobre um determinado assunto e não em recuperar dados que satisfazem sua expressão de busca, nem tampouco documentos, embora seja nestes que a informação estará registrada. Essa característica é o que diferencia os sistemas de recuperação de informação dos Sistemas Gerenciadores de Bancos de Dados (ou simplesmente "bancos de dados"), estudados e implementados desde o nascimento da Ciência da Computação.

Nesse ínterim, cabe mencionar que existem diferentes técnicas no processo de recuperação da informação, conhecidas também como modelos ou modelagem:

> Modelagem em RI é um processo complexo que tem o objetivo de produzir uma função de ranqueamento, ou seja, uma função que atribui escores a documentação em relação a uma consulta. Esse processo pode ser divido em duas tarefas principais: (a) a concepção de um arcabouço lógico para representar documentos e consultas, e (b) a definição de uma função de ranqueamento que computa o grau de similaridade de cada documento em relação a consulta dada. (BAEZA-YATES e RIBEIRO-NETO, 2013, pg. 21).

De acordo com Ferneda (2003), "a eficiência de um sistema de recuperação de informação está diretamente ligada ao modelo que o mesmo utiliza". De acordo com o autor, existem dois tipos de modelos de recuperação da informação, sendo eles o quantitativo e o dinâmico.

Os modelos quantitativos, também conhecidos como modelos clássicos (BAEZA-YATES e RIBEIRO-NETO, 2013, pg. 21), são técnicas de recuperação da informação que se fundamenta em disciplinas como lógica, estatística e teoria dos conjuntos. Muitos desses modelos foram criados na década de 1960 e 1970, porém suas principais ideias são utilizadas até hoje em mecanismos de busca na web. Aqui estão incluídos: modelo booleanos, modelo vetorial e o modelo probabilístico (BAEZA-YATES e RIBEIRO-NETO, 2013, pg. 24).

Já os modelos dinâmicos reúnem métodos mais recentes de recuperação da informação que utilizam técnicas oriundas da inteligência artificial e que representam alternativas promissoras a serem desenvolvidas e estudadas mais profundamente. Esses modelos permitem que o usuário interaja e interfira diretamente no processo de recuperação da informação, já que, diferente dos modelos quantitativos, o modelo dinâmico permite uma evolução e adaptação da

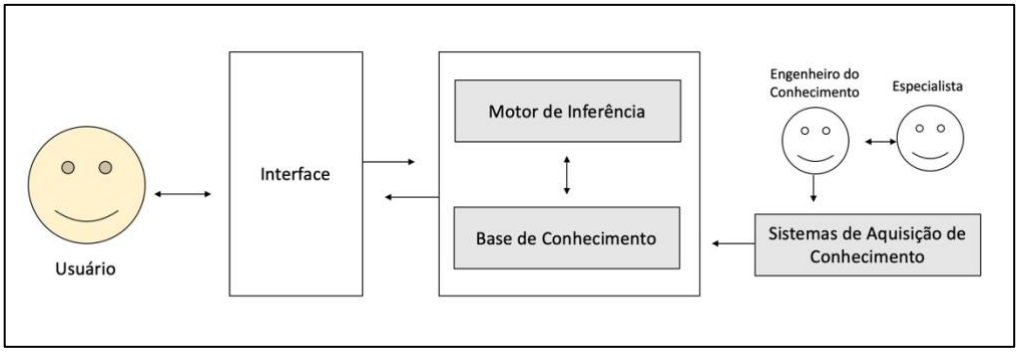
informação de acordo com os interesses de busca do usuário e interações com o sistema (FERNEDA, 2003).

Com isso, para este trabalho, optou-se por analisar a recuperação da informação com base nos modelos dinâmicos, já que uma de suas técnicas utilizadas, os sistemas especialistas, é a base fundamental de um *chatbot*. Sistemas especialistas podem ser definidos como:

> [...] um sistema computacional que procura representar o conhecimento de um especialista humano em um domínio particular, de maneira a auxiliar na tomada de decisões e na resolução de problemas relacionados a esse domínio. A ideia subjacente à construção dos sistemas especialistas é que a inteligência não é apenas raciocínio, mas também memória. É comum considerarmos inteligente uma pessoa que possui grande quantidade de informação sobre um determinado assunto. Assim, os sistemas especialistas obedecem ao princípio de que memória é condição necessária para a inteligência (FERNEDA, 2003).

Tratam-se de sistemas de recuperação desenvolvidos com o objetivo de serem especialistas de um determinado domínio (assunto). Normalmente, são sistemas baseados em perguntas e respostas que, por meio da interação do usuário, sugere uma solução para o problema em questão. Os sistemas especialistas são construídos e arquitetados com base nos seguintes pilares: sistema de aquisição de conhecimento, base de conhecimento, motor de inferência e a interface com o usuário.

Ilustração 3: Estrutura de um sistema especialista



Fonte: Adaptado de Ferneda (2003)

Interface: É a plataforma utilizada para estabelecer a comunicação entre o usuário e o sistema especialista, sendo que a interação pode ter um formato de perguntas e respostas (diálogo) ou em um formato de questionário com perguntas pré-fixadas. Embasado na interação do usuário, o sistema é capaz de inferir e propor uma solução final de acordo com a base do conhecimento previamente alimentada. (FERNEDA, 2003).

Motor de Inferência: Trata-se do intermediador entre a base de conhecimento e o que o usuário pergunta/responde. Então, pode se dizer que o motor de inferência analisa a informação fornecida pelo usuário, busca e infere a melhor resposta dentro da base de conhecimento do sistema e retorna uma possível solução para o problema inicial do usuário (FERNEDA, 2003).

Base de Conhecimento: Pode se dizer que esse é o coração de um sistema especialista. Onde está toda a base de conhecimento que fornece as possíveis soluções para os problemas dos usuários. Esse conhecimento geralmente está armazenado e representado no formato de condição-ação: SE "condição" ENTÃO "Ação" (FERNEDA, 2003). Exemplo: SE "gripe" ENTÃO "Dorflex".

Sistemas de Aquisição do Conhecimento: Trata-se de sistemas que auxiliam o Engenheiro do Conhecimento e o especialista humano na representação e indexação do conhecimento dentro da base de conhecimento. Nesse caso, o engenheiro do conhecimento é o profissional responsável por interpretar e extrair, dos especialistas humanos, o conhecimento para representá-lo na base do conhecimento. Logo, o especialista humano é justamente a pessoa que detém o conhecimento que será utilizado e "emulado" pelos sistemas especialistas.

Em suma, BAEZA-YATES e RIBEIRO-NETO (2013) observam que a recuperação da informação pode ser estudada sob 2 aspectos, conforme detalhados a seguir:

> Em termos de pesquisa, a área pode ser estudada sob dois pontos de vista distintos e complementares: um centrado no computador e outro centrado no usuário. Na visão centrada no computador, a RI consiste principalmente na construção de índices eficientes, no processamento de consultas com alto desempenho e no desenvolvimento de algoritmos de ranqueamento, a fim de melhorar os resultados. Na visão centrada no usuário, a RI consistente principalmente em estudar o comportamento do usuário, entender suas principais necessidades e determinar como esse entendimento afeta a organização e operação do sistema de recuperação.

Dessa forma, com base na citação, este trabalho terá como foco principal estudar a recuperação da informação sob o viés centrado no usuário. Usuário aqui entende-se tanto o usuário que irá utilizar os *chatbots* para a recuperação da informação, como o indivíduo responsável por desenvolvedor esses sistemas. Por meio do viés centrado no usuário, é possível entender melhor como os *chatbots* estão contribuindo para a recuperação da informação e, principalmente, como a CI pode colaborar nesse processo. Em suma, a eficiência de um

sistema de recuperação da informação é reflexo de como a informação foi organizada para ser encontrada. E, nesse sentido, os Sistemas de Organização do Conhecimento empregam um papel importante nesse processo.

2.4. Sistemas de Organização do Conhecimento

Antes de discorrer sobre o que são Sistemas de Organização do Conhecimento, convém contextualizá-los no âmbito da Organização e Representação da Informação e do Conhecimento. Desde os primórdios, a civilização humana sempre utilizou de artefatos, técnicas e símbolos para representar a realidade (SCHIESSL; SHINTAKU, 2012). Um exemplo disso ocorre quando o homem começou a pintar figuras rupestres em cavernas com o intuito de representar alguma atividade ou prática do seu dia a dia. Nesse sentido, a palavra "representar" significa colocar algo no lugar de alguma coisa que, por conseguinte, também está atrelado ao conceito de substituição (ALVERENGA, 2003 apud SCHIESSL; SHINTAKU, 2012).

> Representar é o ato de utilizar elementos simbólicos – palavras, figuras, imagens, desenhos, mímicas, esquemas, entre outros – para substituir um objeto, uma ideia ou um fato. (SCHIESSL; SHINTAKU, 2012)

Na Ciência da Informação, a palavra representação está associada com o fato de simbolizar a informação ou conhecimento (SCHIESSL; SHINTAKU, 2012). De forma bem resumida, trata-se da simbolização do conhecimento/informação com o intuito de recuperação. Isso é feito utilizando uma linguagem específica, denominada de linguagem documentária ou linguagem de indexação, que tem o intuito de representar os conceitos chave de um documento para que o mesmo possa ser recuperado a partir da elaboração de estratégias de busca em um sistema de recuperação da informação que utilize essa linguagem.

> Entende-se, então, que a linguagem documentária visa representar, de forma padronizada, o conhecimento de domínio específico e estabelecer a ligação entre o usuário e o conteúdo do documento. Ela é construída artificialmente com o propósito de sintetizar por meio de vocabulário controlado, o conteúdo de documentos, e ser utilizada em sistemas de indexação, armazenamento e recuperação de informação. Enfim, delimita o vocabulário de determinado domínio de interesse de forma clara e precisa. (SCHIESSL; SHINTAKU, 2012)

A linguagem documentaria é uma linguagem utilizada na Ciência da Informação com objetivo de traduzir a linguagem natural (a linguagem falada no cotidiano que emprega emoções, características culturais etc.) para uma linguagem padronizada e estável para, posteriormente, facilitar a recuperação da informação, seja pelas pessoas ou por sistemas de recuperação de

informação. A linguagem de indexação é construída de forma artificial com o intuito de sintetizar um conteúdo para uma linguagem mais clara e objetiva, facilitando a compreensão dos sistemas e dos usuários no momento da recuperação da informação. Atualmente, no contexto da Ciência da Informação, pode se dizer que os Sistemas de Organização do Conhecimento são uma denominação atual para as linguagens documentárias, pois esse último elemento agrega elementos incorporados das inovações tecnológicas da era digital (CARLAN, 2011).

Os Sistemas de Organização do Conhecimento (SOCs) - do inglês *Knowledge Organization System*, ou simplesmente KOS – são instrumentos que tratam das representações do conhecimento, utilizando modelos de abstração do mundo e construídos para determinada finalidade, como a de facilitar a recuperação da informação (SCHIESSL; SHINTAKU, 2012).

Carlan (2011) define SOC como:

> No contexto da Ciência da Informação, os SOC ou esquemas de representação do conhecimento são instrumentos que fazem a tradução dos conteúdos dos documentos originais e completos, para um esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, consequentemente, facilitar a recuperação das informações contidas nos documentos.

Como complemento, Schiessl e Shintaku (2012) agregam que SOCs são:

> Sistemas conceituais semanticamente estruturados que contemplam termos, definições, relacionamentos e propriedades dos conceitos. Na organização e recuperação da informação, os SOC cumprem o objetivo de padronização terminológica para facilitar e orientar a indexação e usuários. Quanto à estrutura variam de um esquema simples até o multidimensional, enquanto suas funções incluem a eliminação da ambiguidade, controle de sinônimos ou equivalente e estabelecimento de relacionamentos semânticos entre conceitos

Já a Enciclopédia de Organização do Conhecimento (MAZZOCCHI, 2014 apud Souza et al.'s, 2012, tradução nossa), traz uma definição ainda mais abrangente sobre SOC:

> Sistema de Organização do Conhecimento (SOC) é um termo genérico usado para se referir a uma ampla gama de itens (por exemplo, títulos de assuntos, tesauros, esquemas de classificação e ontologias), que foram concebidos com relação a diferentes finalidades, em momentos históricos distintos. Eles são caracterizados por diferentes estruturas e funções específicas, maneiras variadas de se relacionar com a tecnologia e usados em uma pluralidade de contextos por diversas comunidades. No entanto, o que todos eles têm em comum é que foram projetados para apoiar a organização de conhecimento e informação, a fim de facilitar o gerenciamento e a recuperação.

Em resumo, observa-se que os SOC são fundamentais para a qualidade de recuperação da

informação, pois eles consistem em técnicas e procedimentos necessários para que um sistema de recuperação da informação possa funcionar de modo a satisfazer a necessidade de informação de um determinado usuário.

De acordo com Vickery (2008, apud SCHIESSL; SHINTAKU, 2012), os SOCs possuem diferentes propósitos, ao passo que os sistemas de informação foram evoluindo:

> Era da pré-ordenação: estruturas estáticas que atendem às necessidades dos sistemas manuais de organização e recuperação da informação, como índices e catálogo;
>
> Era da pós-coordenação: mais dinâmicos, possibilitam que cada um de seus elementos sejam manipulados de forma independente para restar os assuntos de cada documento, como vocabulários controlados e tesauros;
>
> Era da Internet: classificações hierárquicas que orientam o usuário na escolha do termo que melhor expressa sua questão de busca como taxonomias;
>
> Era da web semântica: São projetados por agentes inteligentes, como ontologias.

No âmbito das funcionalidades dos SOCs, González (2011) lista algumas:

> Permitem categorizar recursos com esquemas de organização determinados, para facilitar sua posterior recuperação;
>
> Permitem ao usuário interpretar estruturas de conhecimento de maneira organizada;
>
> Contribuem para melhorar a eficiência dos serviços de informação;
>
> Buscam a terminologia especifica das instituições (a maioria da informação é de origem interna);
>
> Muitas vezes, provêm do trabalho cooperativo no ambiento empresarial (extranets e comunidades de trabalho);

A construção e a eficiência de qualquer SOC parte do desenvolvimento de sua base fundamental: os vocabulários controlados. Vocabulário controlado é "[...] uma coleção de termos que possibilitam descrever um documento com maior precisão. A meta principal é garantir a descrição mais objetiva e facilitar a recuperação dos conteúdos que integram um SOC" (COLOMBO, 2002 apud SCHIESSL; SHINTAKU, 2012). Portanto, são essenciais para a recuperação da informação, já que a linguagem natural carrega características linguísticas próprias de cada cultura, que podem interferir na descrição e, por conseguinte, na recuperação da informação.

Alguns exemplos dessas características linguísticas são a sinonímia, a polissemia e a homografia. A primeira é a sinonímia, que consiste na relação de sentido entre as palavras com

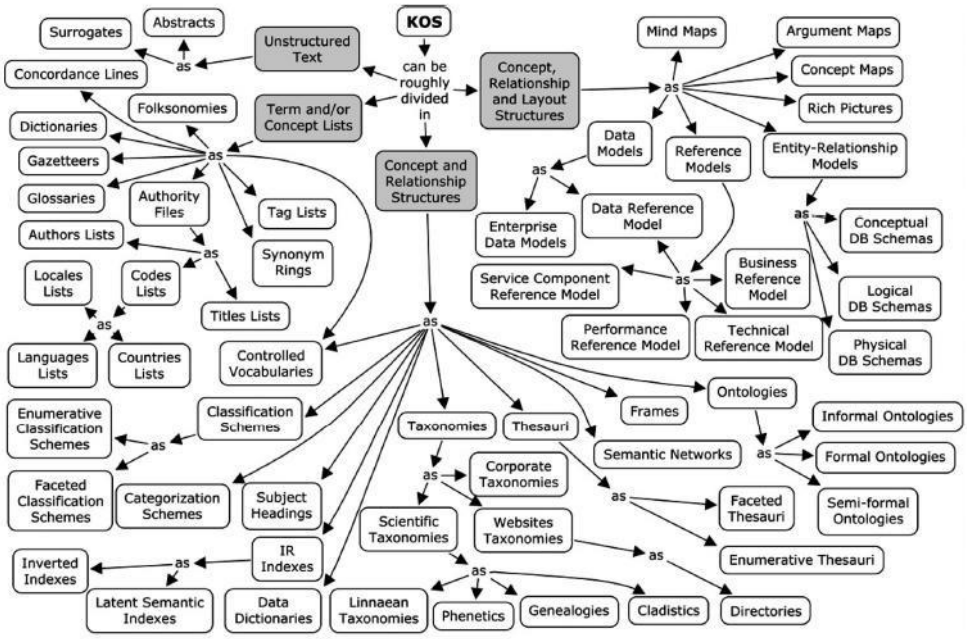
significados semelhantes, isto é, mais de um termo representando o mesmo conceito. Já a polissemia, é o episódio onde um termo possui mais de um significado ou aceitação. Por fim, a homografia, que são palavras que possuem a mesma escrita, porém com significados diferentes. Um clássico exemplo é a palavra manga, que pode significar uma fruta, mas também parte de uma camisa. (SCHIESSL; SHINTAKU, 2012).

No contexto de organização do conhecimento e, principalmente em relação ao controle de vocabulário, se convém distinguir "termos" e "palavras", já que são conceitos importantes para o desenvolvimento de linguagens artificiais:

> Termo é uma palavra especializada, ou o conjunto delas, isto é, ele está inserido no contexto das línguas especializadas. As palavras, na língua natural. A palavra pode ter vários significados e o termo possui apenas um significado. Isso porque o termo se relaciona univocamente com o conceito especializado em uma área do conhecimento. A palavra não está restrita ao domínio, mas ao fenômeno da comunicação pela língua natural. (SCHIESSL; SHINTAKU, 2012)

Dessa forma, os vocabulários controlados são utilizados para representar um conteúdo abrangente de um domínio específico em uma linguagem artificial, utilizando termos que são definidos de forma não ambígua e que não tenha redundâncias. Portanto, cria-se uma linguagem artificial própria, que tem por objetivo mitigar problemas, como de sinonímia, polissemia e homografia. O resultado do uso de um vocabulário controlado são SOCs mais eficientes, atendendo as reais necessidades de informação de um usuário. Existem diversos Sistemas de Organização do Conhecimento, conforme Ilustração 4.

Ilustração 4: Sistemas de Organização do Conhecimento



Fonte: Mazzocchi (2014 apud Souza et al.'s, 2012)

Esses sistemas possuem objetivos similares em relação a organização do conhecimento. Entretanto, cada um deles é característico para cada situação. Existem os sistemas específicos para o controle de textos não estruturado, como resumos e linhas de concordâncias. Existem aqueles que se preocupam em organizar listas de conceitos e termos, como folksonomias, dicionários e glossários. Além disso, existem os sistemas focados na visualização do conhecimento, tais como mapas mentais, modelos de referências, modelos de dados, entre outros. Por fim, os sistemas que tratam do relacionamento entre os termos, como as taxonomias, tesauros e ontologias. Esses últimos serão o foco desta pesquisa como sugestão para melhoria do desempenho de um *chatbot*.

Em suma, desde o surgimento da escrita e da comunicação, o homem sempre esteve procurando maneiras de organizar o conhecimento para o seu acesso. Com o advento das tecnologias da informação e comunicação, especialmente com o surgimento da internet, essa preocupação se tornou mais evidente e, por que não, urgente. Com a miríade de informações geradas constantemente, se faz necessário o uso de técnicas para organizá-las e, em consequência recuperá-las. Assim, a Ciência da Informação, mais especificamente, a Recuperação da Informação, possibilitam uma reflexão social sobre o uso da informação e, principalmente, como essas informações podem ser organizadas para serem encontradas de forma simples pelas pessoas.

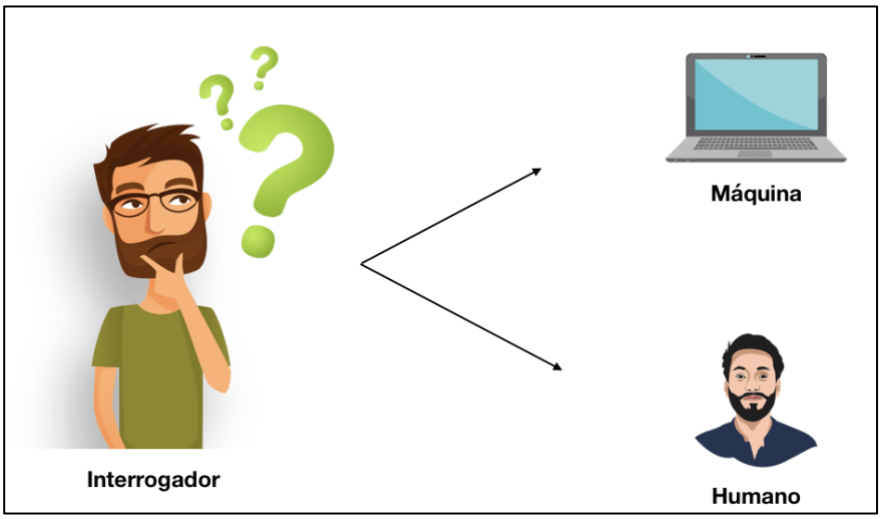
3. BEM-VINDO À ERA DOS CHATBOTS

Em 1950, o matemático britânico Alan Turing se questionou sobre a possibilidade de máquinas raciocinarem: "*Can machines think?* (Podem as máquinas pensar?)". A indagação levantada por Turing originou um teste que foi chamado de "Jogo da Imitação", também conhecido como "Teste de Turing" (LEONHARDT, 2005). Esse teste foi proposto para determinar se um ser humano seria capaz de conversar com uma máquina sem perceber.

> O jogo da imitação foi proposto como forma de caracterizar e estudar a possibilidade de se produzir uma máquina dita pensante. Na sistemática original do jogo, três pessoas são envolvidas: um homem (A), uma mulher (B) e um interrogador (C). O interrogador pode ser de qualquer sexo. Ele permanece num quarto, separado dos outros dois. Assim, o objetivo do jogo, para o interrogador, é descobrir quem é o homem e quem é a mulher (LEONHARDT, 2005).

Resumidamente, o objetivo do jogo era que o investigador (C) identificasse quem era o homem (A) e quem era a mulher (B). As perguntas eram feitas de forma datilografada, ou seja, o investigador não poderia ver os personagens e, obviamente, ouvir suas vozes. Usando essa mesma dinâmica, Turing propôs um teste semelhante, tendo agora como personagens um humano e uma máquina. Com isso, o interrogador deveria adivinhar, por meio de perguntas também datilografadas (Ilustração 5), se estava se comunicando com um ser humano ou com alguma espécie de inteligência artificial. Caso a máquina fosse confundida com um humano, essa passaria no teste e seria classificada como inteligente.

Ilustração 5: Esquematização do Teste de Turing



Fonte: O autor (2020)

Alencar, Schmitz e Cruz (2018) indagam que a questão hoje em dia não é mais se um ser humano é capaz de perceber que está falando ou não com um computador, mas sim por quanto tempo ele pode ser "enganado" pela máquina. A verdade é que, desde que se criou o teste proposto por Turing, surgiram diversas pesquisas e ideias com o objetivo de desenvolver aplicações que pudessem promover uma comunicação em linguagem natural entre o usuário e o computador (TURING, 2009). Os *chatbots* nascem nesse contexto como proposta de agente inteligente para interação com o ser humano.

3.1. Sobre os *chatbots*: definição, história e evolução

Chatbots (ou *Chattterbots*) é uma categoria de robôs (ou agentes inteligentes) que simulam a conversa com um ser humano por meio de uma máquina (LEONHARDT, 2005). É um tipo de plataforma conversacional que tem por objetivo ajudar o ser humano em diversas atividades. O termo *chatbot* é a junção das palavras em inglês: *Chat* (conversa) e *Bot* (robô). A seguir, algumas definições, de diferentes autores, em relação ao termo *chatbot*:

O pesquisador Laven (2020) traz uma definição bastante objetiva sobre o que é um *chatbot*:

> O Chatterbot é um programa que tenta simular conversas digitadas, com o objetivo de enganar temporariamente um ser humano a pensar que estava falando com outra pessoa. (tradução do autor)

Com um pouco mais de detalhes, Muldowney (2017) descreve *chatbots* como:

> Um *chatbot* é um programa de computador criado para participar de uma conversa. Normalmente, os *chatbots* são escritos para interagir com seres humanos (em vez de outros *chatbots*) e o fazem por uma variedade extremamente ampla de razões. Nos negócios, eles estão proliferando como uma alternativa aos sites: em vez de um cliente ter que tomar a iniciativa pesquisando nas páginas do site, o *chatbot* fornece ao cliente um guia interativo, que pode orientá-lo em relação ao produto que estava procurando e até organizar pagamentos e entrega. Da mesma forma, as organizações que fornecem uma grande quantidade de informações on-line para clientes (como organizações de saúde ou órgãos governamentais) usam *chatbots* para ajudar os clientes a obterem as informações desejadas por meio de uma conversa em vez de um mecanismo de pesquisa. (tradução do autor)

De maneira análoga, Leonhardt (2005) diz que *chatbots* são:

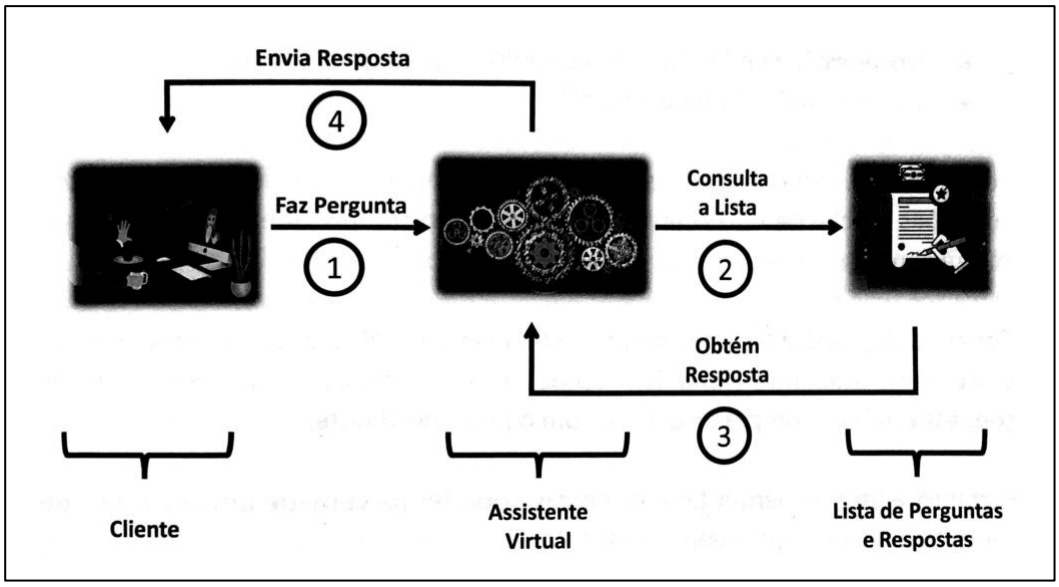
> [...] sistemas que procuram simular uma conversação similar à de um ser humano a fim de enganar, ao menos momentaneamente, um interlocutor. Essa característica desperta o interesse da comunidade científica e industrial, pois, além desses sistemas proverem uma interface mais amigável e natural, eles permitem explorar as relações sociais que os usuários desenvolvem com máquinas.

Por fim, a Gartner (2020) – uma importante empresa americana em pesquisa e consultoria em tecnologia - apresenta *chatbot* como:

> [...] uma interface conversacional específica do domínio que usa um aplicativo, plataforma de mensagens, rede social ou solução de bate-papo para suas conversas. Os *chatbots* variam em sofisticação, desde truques simples de marketing baseados em árvore de decisão até implementações construídas em plataformas ricas em recursos. Eles sempre têm escopo estreito. Um *chatbot* pode ser baseado em texto ou voz ou uma combinação de ambos.

Mediante o exposto, um *chatbot* é um programa de computador que tem por objetivo responder as perguntas dos usuários de tal maneira que o indivíduo tenha a impressão de estar conversando com uma pessoa real. Obviamente que, ao interagir com esses agentes, nota-se que a comunicação está sendo intermediada por um robô. Logo, a expressão utilizada por alguns autores em relação a tentar enganar o usuário, parece não fazer muito sentido nos dias atuais, visto que a interação com os *chatbots* tem tornado rotina e o objetivo deste recurso é facilitar a busca pela informação. O fato é que esses agentes são alimentados por informações oriundas de diversos repositórios, tais como: livros, bancos de dados, acervos digitais, websites, documentos institucionais. São informações "estáticas" que servem de insumos para compor a base de conhecimento desses sistemas, e, consequentemente, serem recuperadas de maneira natural por meio de conversas. Em outras palavras, um *chatbot* é um modo de, literalmente, bater um papo com a informação.

Ilustração 6: Esquematização do funcionamento de um *chatbot*



Fonte: Alencar, Schmitz e Cruz (2018)

Desde a década de 1960, as tentativas de se criar um *chatbot* foram diversas (LAVEN, 2020). Ao longo da história, Rothermel e de Souza Domingues (2007) identificam três gerações de *chatbots*. A primeira geração baseia-se em técnicas que combinam padrões e regras gramaticais. A segunda, traz técnicas de inteligência artificial, como as regras de produção de redes neurais. Por fim, a última e mais recente geração de *chatbots* está baseada na utilização de linguagens de marcação para a construção de suas bases de conhecimento, como AIML (*Artificial Intelligence Markup Language*). Essas três gerações podem ser representadas, respectivamente, pelos *chatbots* Eliza, Julia e A.L.I.C.E.

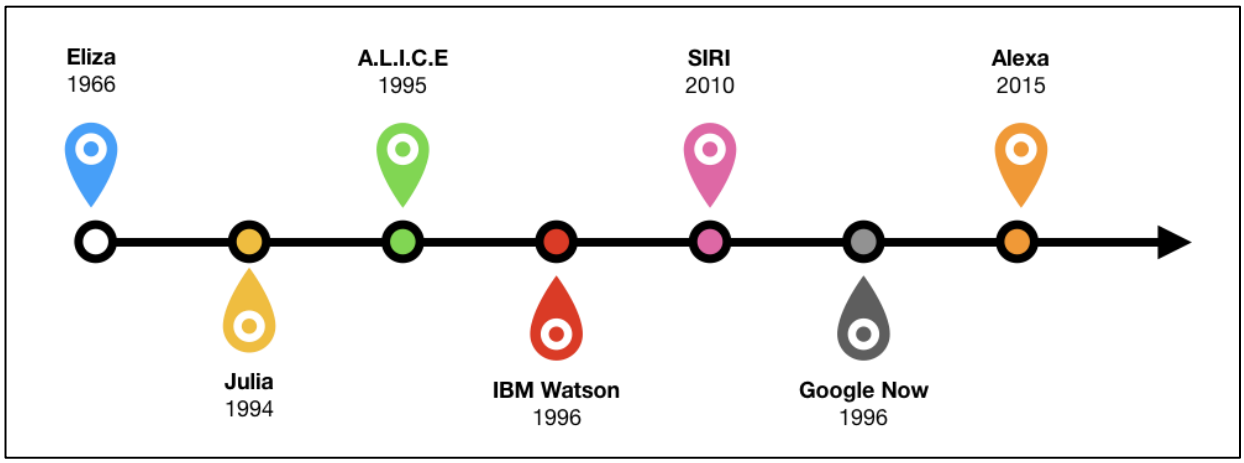
Eliza foi o primeiro *chatbot* criado no ano de 1966 por Joseph Weizenbaum, no Massachusetts Institute of Technology (MIT). Na ocasião, o objetivo do agente era simular um psicanalista rogeriano em uma conversa com um paciente, interagindo com ele por meio de perguntas sobre as questões feitas pelo próprio usuário (COMARELLA, 2008). Era um *chatbot* pequeno, com cerca de 200 linhas construído sob regras compostas por palavras-chave e um interpretador de dados digitados pelo indivíduo. Ao digitar uma frase de entrada, o interpretador fazia uma seleção das palavras-chave dessa frase e comparava com o conhecimento do *chatbot.* Logo em seguida, uma resposta era dada ao usuário utilizando um novo conjunto de palavras de forma estruturada, no formato de frase. Apesar de toda a interação, o agente não era capaz de contextualizar uma conversa ou sequer lembrar o que falou anteriormente (LAVEN, 1996 apid COMARELLA, 2008).

Em 1994, Michael Maudin desenvolveu o *chatbot* Julia, que diferente da Eliza, foi criada utilizando princípios da Inteligência Artificial, como a técnica de redes neurais que busca simular o cérebro humano para dialogar (MOURA, 2003). Na ocasião, o *chatbot* foi desenvolvido para ser um assistente de um jogo de realidade virtual. Julia possuía internamente variáveis contextuais que poderiam ser atualizadas no momento em que a conversa estava se desenvolvendo. Ela era capaz de contextualizar o diálogo, gerando maior sensação de realidade (MOURA, 2003).

Por fim, em 1995, o pesquisador Richard Wallace criou A.L.I.C.E (*Artificial Linguistic Internet Computer Entity*). Trata-se do primeiro *chatbot* criado com base na linguagem AIML na qual é utilizada atualmente em alguns desses agentes para especificar bases de conhecimento. Diferente dos outros *chatbots*, A.L.I.C.E. operava como um modelo de aprendizagem

supervisada, ou seja, era necessário que um especialista monitorasse suas conversas com o objetivo de aperfeiçoar suas respostas futuras. (MOURA, 2003).

Ilustração 7: Linha do tempo dos chatbots



Fonte: O autor (2020)

Conforme a Ilustração 7, o desenvolvimento desses *chatbots*, principalmente com o surgimento da A.L.I.C.E., incentivou, nos últimos anos, a criação de diversos outros sistemas, com diferentes finalidades, como o Watson da IBM, a Siri da Apple, Google Assistant/Google Now e a Alexa da Amazon. Esses agentes, vem ganhando força em diversos segmentos do mercado nos últimos anos e utilizam interfaces agradáveis e podem ser incorporadas em diversos canais e plataformas de mensagem instantânea (ou IM – *Instant Messaging*), como: Facebook Messenger, WhatsApp, Telegram, Slack, Skype. Não obstante, os *chatbots* podem ser integrados a sistemas telefônicos, atendendo e auxiliando o usuário antes de transferi-los para um ser humano. Além disso, podem ser incorporados em sistemas de e-mail, onde muitas das vezes respondem as mensagens eletrônicas realizando tarefas de uma secretária executiva, como agendar reuniões (EXCEED.AI, 2020).

Atualmente, os *chatbots* são utilizados para diferentes finalidades (ALENCAR, SCHMITZ E CRUZ, 2018), tais como: responder perguntas frequentes de um produto ou serviço; outros com o propósito de atuarem como um sistema de entretenimento; existem também os *chatbots* que foram criados com o propósito de captação de contatos para estímulos de marketing e vendas; e os que servem para fornecer informações sobre saúde e bem-estar. Além disso, existem *chatbots* com funções de serem Assistentes Virtuais, como o Google Assistant, a Alexa, a Siri, entre outros.

É importante salientar que Assistentes Virtuais (também conhecidos como Assistentes Virtuais Inteligentes) é um termo utilizado quase que sinonimamente para referir-se a *chatbots*. Alguns autores, entretanto, argumentam que existem diferenças entre os dois termos. Uns acreditam que os Assistentes Virtuais são mais robustos no sentido de integrarem com sistemas corporativos e terem maior capacidade de aprendizagem (ALENCAR, SCHMITZ E CRUZ, 2018). Outros autores argumentam que a principal diferença é em relação as plataformas nas quais esses sistemas são utilizados. Muldowney (2017) versa que enquanto os Assistentes Virtuais vivem em dispositivos domésticos e dispositivos próprios (como smartphones ou caixas de som inteligentes) os *chatbots* habitam amplamente as páginas na internet, bem como aplicativos como WhatsApp e Facebook.

Nesse contexto, Dale (2016) simplifica o embate entre Assistentes Virtuais e *Chatbots* com o seguinte argumento:

> Para os propósitos atuais, usaremos o termo "*chatbot*" para nos referir a qualquer aplicativo de software que entre em diálogo com um humano usando linguagem natural. O termo é mais frequentemente usado em conexão com aplicativos que conversam via linguagem escrita, mas com avanços no reconhecimento de fala, que cada vez mais parece um diferenciador bastante falso.

Em síntese, cada autor traz pontos pertinentes entre essas diferenças. Por exemplo, os *chatbots*, de fato, parecem estar mais focados em um contexto conversacional bastante específico, enquanto os Assistentes Virtuais estão focados em um domínio de conversa mais amplo, justamente por se comportar como um assistente pessoal. Esta pesquisa entenderá os Assistentes Virtuais como um tipo de *chatbot* e, portanto, o termo *chatbot* que será utilizado aqui partilha o argumento de Dale (2016), que o enxerga como qualquer aplicativo que entre em diálogo com um humano por meio de linguagem natural. De certo, os *chatbots* estão cada vez mais precisos, inteligentes e humanos, graças ao uso e a incorporação de tecnologias oriundas da Inteligência Artificial.

3.2. O uso de Inteligência Artificial nos *chatbots*

Inteligência Artificial (IA) é um grande e complexo campo de estudos da Ciência da Computação que possui como propósito estudar técnicas de programar computadores para simular a inteligência humana.

GABRIEL (2018), eu seu livro "*Você, Eu e os Robôs*", define inteligência artificial como:

> Inteligência Artificial é um termo amplo, que abriga debaixo de seu guarda-chuva as inúmeras disciplinas envolvidas com o desenvolvimento de máquinas pensantes. Como o objetivo de AI é o desenvolvimento da inteligência, ela se relaciona com todas as áreas do conhecimento que são usadas no estudo da inteligência, além de métodos, algoritmos e técnicas que possam tornar um software/hardware inteligente no sentido humano da palavra.

John McCarthy (2007, tradução nossa), considerado o pai da Inteligência Artificial, define o termo como:

> É a ciência e a engenharia de fabricação de máquinas inteligentes, especialmente programas de computador inteligentes. Está relacionado à tarefa semelhante de usar computadores para entender a inteligência humana, mas a IA não precisa se limitar a métodos que são biologicamente observáveis.

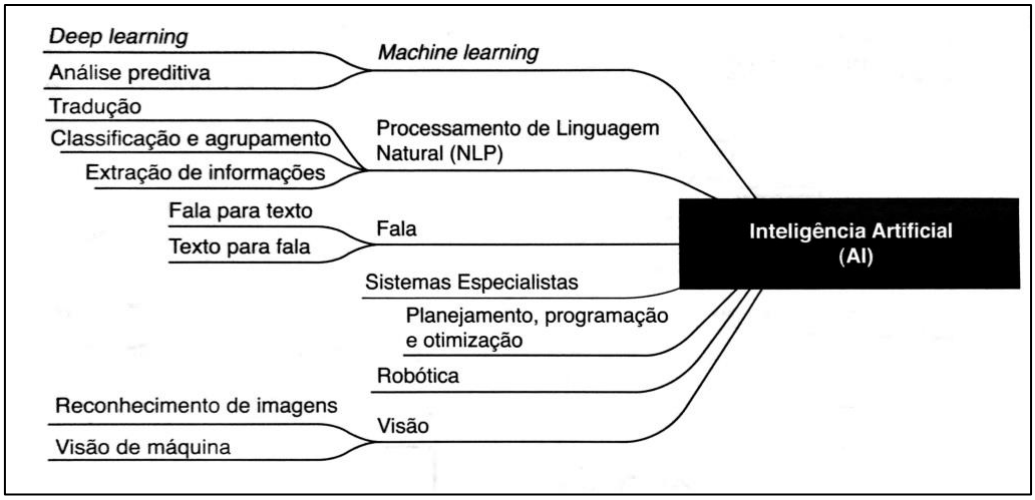
Russell e Norving (2004, tradução nossa) autores do livro Inteligência Artificial, considerado um trabalho referencial na literatura sobre inteligência artificial, define-a como:

> [...] é o desenho e desenvolvimento de agentes inteligentes que recebem percepções a partir do meio ambiente para realizar ações que afetam esse ambiente.

Como dito, o campo da Inteligência Artificial é bastante complexo de ser definido. Apesar das sutis diferenças em cada uma das definições, observa-se que todas as definições concordam que a Inteligência Artificial é uma área que estuda maneiras de fazer com que máquinas simulem a inteligência humana. Nesse sentido, tal como a inteligência humana, a inteligência artificial é uma área considerada multi-inter-trans-disciplinar (GABRIEL, 2018), e, por conseguinte, se conecta a diversas disciplinas, tais como a Filosofia, Sociologia, Matemática, Biologia, Neurociência, Psicologia, Ciência da Computação, Linguística, Ciência da Informação.

O termo Inteligência Artificial foi cunhado pela primeira vez em 1956 por um grupo de pesquisadores liderados pelo cientista John MacCarthy (RUSSELL E NORVING, 2004), que tinham como objetivo descrever um mundo em que as máquinas seriam capazes de executar trabalhos realizados por humanos. Entretanto, apesar do termo ter sido cunhado oficialmente nos anos 50, a presença de homens mecânicos e seres artificiais fazem parte do imaginário mitológico humano desde a Grécia Antiga (GABRIEL, 2018), onde já se tinha a ideia da criação de seres autônomos capazes de realizar tarefas básicas, tais como limpar, servir comida, proteger.

Figura 3: Disciplinas da Inteligência Artificial



Fonte: Gabriel (2018)

Conforme a Figura 3, a Inteligência Artificial vincula-se a diversas subáreas que são estudadas e que compõem o grande campo da AI. Semelhante as habilidades humanas, essas subáreas são dividias em estudos relacionados de aprendizagem, interpretação, fala e comunicação, visão, robótica, automação, entre outras. Para simular essas habilidades de forma artificial, diversos algoritmos e modelos matemáticos são utilizados para que essas máquinas possam chegar o mais próximo possível do pensamento e inteligência humana.

> Note-se que a conexão da IA com a robótica tem o mesmo tipo de relação entre a mente e o corpo humano: nossa mente precisa do corpo (incluindo o cérebro físico) para "sentir" (visão, olfato, paladar, tato e audição), "processar" (pensamento) e "atuar" no mundo, e todas essas funções acontecem em conjunto entre o corpo e a mente para: perceber, agir, receber *feedback*, e ajustar a próxima ação, no caso da IA, ocorre o mesmo processo entre software (mente) e hardware (corpo), de forma que são interdependentes e acionados em diferentes graus de intensidade, de acordo com a função que se exerce em cada momento. (GABRIEL, 2018)

Atualmente, a Inteligência Artificial vem sendo utilizada por diversas empresas e aplicada no desenvolvimento e aprimoramento de distintas ferramentas, como o *chatbot*. Os *chatbots* estão cada vez mais eficientes, devido a utilização de técnicas avançadas que compõe o campo da Inteligência Artificial e que simulam as habilidades humanas, tais como: *Machine Learning, Deep Learning, Natural Language Processing, Text Analytics/Text Mining, Speech Recognition, Image Recognition* e *Emotion Recognition*.

Machine Learning

Machine Learning (ML), também conhecida como "aprendizagem de máquinas", é uma das

subáreas da Inteligência Artificial que utiliza algoritmos que permitem que um programa aprenda sem que seja expressamente programado por um ser humano (GABRIEL, 2018). Estes algoritmos são capazes de reconhecer padrões e similaridades baseados em suas experiências anteriores. Por meio de modelos matemáticos, sobretudo aqueles que simulam as redes neurais do cérebro humano e que são estudados dentro de outra subárea - *Deep Learning (DP)* -, os seres humanos podem treiná-los para que sejam capazes de reconhecer imagens, comportamentos, linguagens etc.

Nesse contexto, as técnicas de *Machine Learning* possibilitam que os *chatbots* sejam mais inteligentes e entendam, de forma mais satisfatória, as reais necessidades informacionais de um usuário. Um exemplo é quando um *chatbot*, acoplado a uma camada de ML e DP, consegue identificar a intenção de pergunta de um usuário por meio da análise da frase que esse indivíduo imputa no sistema. Dessa forma, o sistema de recuperação de informação do *chatbot* fica cada vez mais apurado e, como resultado, tende a recuperar a informação de forma mais precisa. Tudo isso só é possível porque os algoritmos de *Machine Learning* foram treinados anteriormente com frases similares e, desta forma, por dedução e por uma pontuação de probabilidade, conseguem identificar que o usuário busca por uma informação que, outrora, foi útil para outro indivíduo.

Natural Language Processing (NLP)

Natural Language Processing, também conhecida como "Processamento de Linguagem Natural", é área dentro da IA responsável por estudar a capacidade que uma máquina tem de entender a linguagem dos seres humano por meio de diversos algoritmos computacionais (LIDDY, 2001). Esses algoritmos analisam os textos de um documento em diferentes níveis, tais como morfológico, sintático, fonológico, léxico, pragmático, discurso e semântico. Desse modo, a NPL pode ser utilizada em diversos contextos, principalmente, para a recuperação da informação. Ao permitir a análise sob a forma natural que os seres humanos se comunicam, os sistemas de recuperação de informação passam a buscar o que é solicitado não mais por meio de palavras chaves específicas, mas sim por meio de uma compreensão mais semântica do que é solicitado. Essa compreensão, tende, portanto, tornar esses sistemas mais assertivos, atendendo as necessidades de informação de um usuário.

Nessa área, fica claro a importância do Processamento de Linguagem Natural para os *chatbots*. Com essa camada de inteligência, os *chatbots* conseguem interpretar e entender o que o

indivíduo de fato busca. A pergunta que é imputada no sistema pelo usuário é convertida em alguma intenção especifica, seja para retornar alguma informação ou realizar um pedido de pizza, por exemplo. Não obstante, a NPL permite diferenciar, em algumas situações, quando um usuário digita uma palavra incorreta ou utiliza sinônimos e/ou expressões especificas de uma região. A NPL é, portanto, uma das principais tecnologias para o bom desempenho de um *chatbot*.

Text Analytic

Text Analytics, *Text Mining*, ou simplesmente Mineração de Texto, é uma técnica de Inteligência Artificial que utiliza o Processamento de Linguagem Natural para entender e encontrar padrões em um determinado documento do texto (PEZZINI, 2016). O principal objetivo dessa técnica é encontrar termos relevantes em um documento e relaciona-los com base na frequência e temática que aparecem no texto. Com essa técnica, é possível que as informações de um texto possam ser classificadas automaticamente nas seguintes categorias: palavras, cores, conceitos/entidades, tópicos específicos, *stopwords* etc.

Nesse sentido, a *Text Analytics* pode ser utilizada em casos que o *chatbot* não tenha uma resposta pré-programada em sua base de conhecimento que responda à pergunta do usuário. Ao compreender o que o usuário deseja (por meio da NPL), o sistema busca essa informação direto em documentos e/ou sites que tenham alguma relação com a intenção de busca do usuário. Portanto, uma vez que a informação desejada é encontrada, extrai-se um pedaço do texto que faça sentido com o que foi perguntado pelo indivíduo e, logo em seguida, esse trecho é retornado para o usuário como resposta à sua pergunta. Apesar das respostas serem trechos de textos extraídos de documentos - isto é, sem tratamento para adaptar-se a personalidade do *chatbot* -, essa técnica é bastante interessante para evitar que o *chatbot* retorne frases como: “Desculpe, mas não encontrei o que você procura”.

Speech Recognition

Speech Recognition, ou Reconhecimento de Fala é também outra técnica de Inteligência Artificial que está muito ligada ao Processamento de Linguagem Natural e *Machine Learning*. É uma disciplina que estuda métodos que possibilitam o reconhecimento e transcrição de linguagem de maneira automática (IBM, 2020). Por meio desses algoritmos, a máquina é capaz de converter a fala humana em texto (*Speech to Text*); e também a conversão de texto em fala (*Text to Speech*). Essa técnica é muito utilizada em diversas situações, tais como a

geração de legendas automáticas em filmes, buscas por meio de comandos de voz, leitura automática de texto e livros digitais.

A *Speech Recognition* adiciona uma camada de interação ainda mais natural entre o usuário e os *chatbots*. Isso se dá porque a *Speech Recognition* pode ser aplicada em agentes que, por exemplo, atendam clientes via telefone com o objetivo de informar e filtrar o usuário para o setor correto de um *callcenter*. Em outras situações, essa técnica permite o aproveito de funções nativas de interação por voz de plataformas de mensagens, como o WhatsApp e o Facebook Messenger. Dessa maneira, basta que o usuário fale o que deseja, utilizando sua própria voz, para que o *chatbot* compreenda o que foi falado e retorne uma resposta, também em formato de voz. Em síntese, a voz/áudio do usuário é transformada em texto para que, consequentemente, esse texto seja processado por algoritmos de Processamento de Linguagem Natural. Após compreendido e encontrada a informação, essa resposta, em formato de texto, é novamente convertida em áudio para, logo em seguida, ser enviada para o usuário.

Image Recognition

Image Recognition ou "Reconhecimento de Imagem" é outro campo de estudo dentro da Inteligência Artificial e também da Visão Computacional que tem por propósito identificar elementos embutidos em uma imagem ou vídeo (MARUTITECH, 2020). Basicamente, são algoritmos que identificam lugares, pessoas, objetos, logotipos em uma ou um conjunto de imagens. Desta forma, logo que um sistema de reconhecimento de imagem recebe um arquivo de foto, esse arquivo é analisado por esses algoritmos e processados com o intuito de encontrar relações com algo que já foi previamente aprendido ou alimentado em seu banco de dados. Em outros termos, é uma técnica que também faz uso de *Machine Learning* e *Deep Learning.*

Ainda não existem muitos *chatbots* que utilizam essa técnica. Entretanto, adicionar uma camada de inteligência de reconhecimento de imagem em um *chatbot* permite que esse agente recupere a informação desejada por meio de uma imagem. Um exemplo seria um agente especialista em detectar doenças de pele. Bastaria que o usuário tirasse uma foto de alguma mancha específica em sua pele para que o sistema inferisse se trata de algo maligno ou benigno. Outro exemplo, seria um *chatbot* especialista em informações sobre candidatos políticos. Bastaria que o usuário enviasse uma foto de determinado candidato para que fosse identificado e, por conseguinte, as informações sobre ele/ela fossem enviadas para o usuário, o que poderia incluir suas propostas, se teve alguma acusação ou se responde à algum

processo.

Emotion Recognition

Por último, a *Emotion Recognition* ou "Reconhecimento de Emoções" – as vezes também referenciada como Análise de Sentimentos. Trata-se de uma disciplina dentro de AI que tem como finalidade identificar as emoções dos seres humanos, seja por meio de imagens/vídeos – com o uso de Reconhecimento de Imagens - ou por meio de áudio/texto - utilizando o Processamento de Linguagem Natural (YAM, 2015). É uma técnica bastante utilizada por marcas para analisar os sentimentos de seus consumidores na internet. Isso é feito com a análise dos comentários em posts da empresa, *tweets* e/ou conteúdo que tenham relação com a marca. Com a utilização de *Machine Learning* e NPL, esses dados são analisados e agrupados em categorias como: raiva, alegre, insatisfeito, triste e mais.

Ao utilizar o reconhecimento de emoções em um *chatbot*, o diálogo com o usuário poderia ser ajustado em tempo real, e ainda detectar que o usuário está com raiva. O *chatbot* tentaria utilizar um tom mais respeitoso, evitando frases com expressões de brincadeiras ou piadas. Da mesma maneira, ao detectar que um cliente está insatisfeito, o agente poderia, automaticamente, transferir a pessoa para um humano, impedindo maiores frustrações ou desentendimentos no momento da interação. Esses são apenas alguns exemplos de como essa camada de inteligência poderia, eventualmente, tornar o diálogo com um *chatbot* ainda mais humanizado e personalizado.

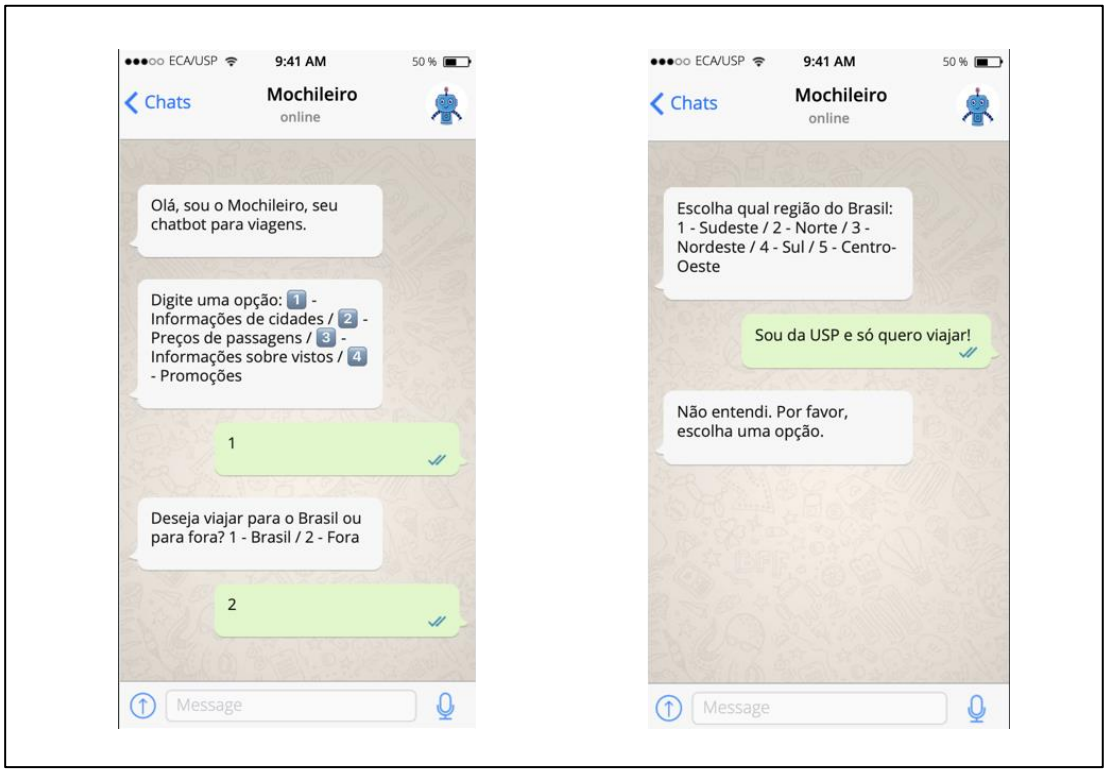
Como visto, a Inteligência Artificial saiu dos filmes para se tornar algo real e aplicável no dia a dia das pessoas. O objetivo aqui não foi tentar esgotar os assuntos levantados sobre AI e suas técnicas. Como versado no começo deste capítulo, a Inteligência Artificial é campo de estudo bastante amplo e que envolve diversas disciplinas. Cada tópico apresentado aqui tem potencial para discursões mais amplas e profundas. Posto isso, o principal objetivo deste subcapítulo foi explicar, de maneira bastante resumida, como cada subárea da Inteligência Artificial pode ampliar a interatividade no momento de recuperar a informação por meio de um *chatbot*. Essas tecnologias, oriundas da Inteligência Artificial, adicionam uma camada de inteligência e dinamismo a esses agentes, permitindo que os usuários encontrem de forma mais precisa o que procuram e, consequentemente, satisfaçam suas necessidades de informação.

3.3. Tipos, elementos e arquitetura dos *chatbots*

De acordo com o relatório técnico da empresa Artificial Solutions (2020), existem dois tipos de *chatbots*: os baseados em regras e os baseados em inteligência artificial.

Os *chatbots* baseados em regras são sistemas que normalmente trabalham com um domínio específico da informação, já que as perguntas e respostas são, em sua grande maioria, previsíveis (ARTIFICIAL SOLUTIONS, 2020). São sistemas que utilizam estruturas baseadas em fluxos de diálogo estruturados, que são acionados de acordo com a palavra-chave encontrada na pergunta do usuário. Com isso, normalmente o desenvolvimento desses *chatbots* constitui em regras como "se/então". Exemplo: Se a pergunta contém a palavra-chave *horário de funcionamento*, então responda com o horário de atendimento do estabelecimento.

Figura 4: Exemplo fictício de *chatbot* de viagens baseado em regras



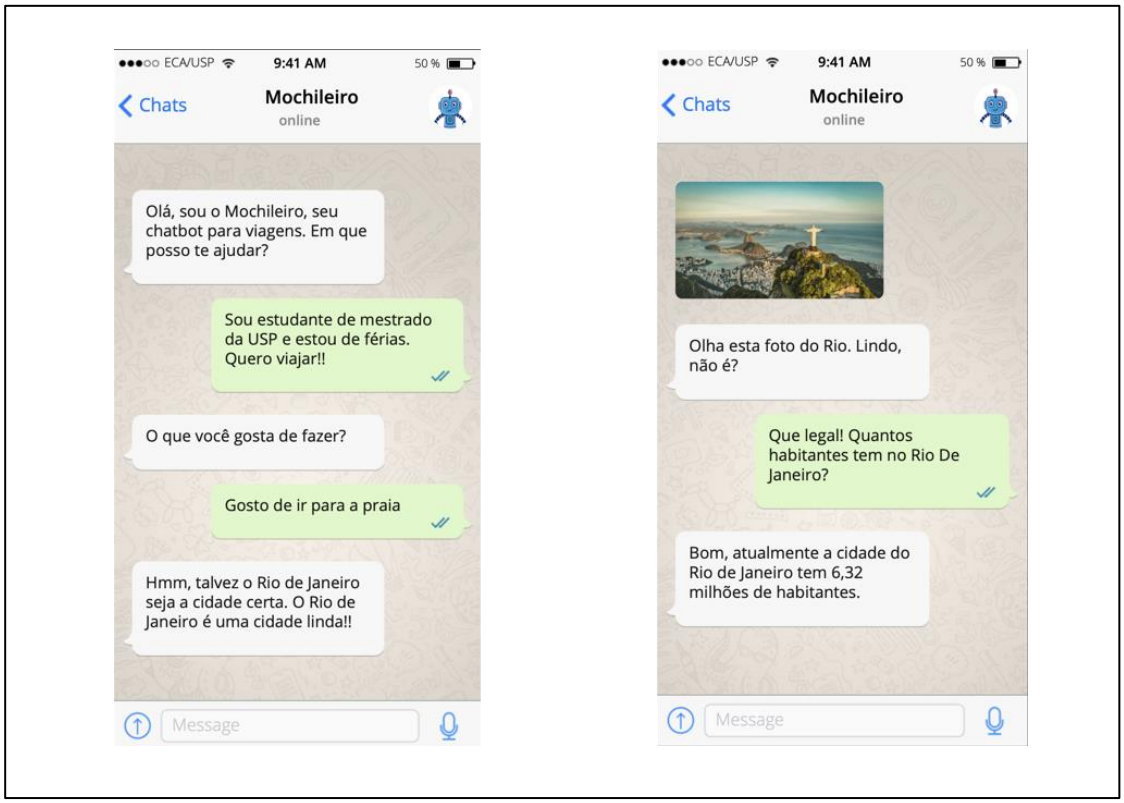
Fonte: O autor (2020)

A Figura 4 exemplifica a interação com um *chatbot* baseado neste tipo de estrutura. Nota-se que todo o diálogo é guiado pela escolha de opções de acordo com a necessidade do usuário. Ao escolher uma opção, a programação do *chatbot* irá processar uma regra: Se opção = 1, então retornar à informação correspondente a essa pergunta. Nota-se que, por ser um agente baseado

em regras com uma estrutura de diálogo bem controlada, ele, normalmente, não reconhece quando o usuário tenta interagir de forma diferente daquela previamente estruturada. Esse é apenas um exemplo de como esses *chatbots* operam. Há também casos nos quais a interação se estabelece por meio de botões. Ao clicar na opção desejada, o usuário é direcionado ao fluxo de diálogo correspondente.

Já os *chatbots* baseados em Inteligência Artificial, são aqueles que utilizam técnicas de AI (como as descritas no tópico anterior) para torná-los mais dinâmicos e precisos em suas respostas (ARTIFICIAL SOLUTIONS, 2020). Uma das características fundamentais desse tipo de agente é a possibilidade de interação por meio de linguagem natural. Com tal característica, basta que o usuário pergunte algo em linguagem natural para que o sistema seja capaz de identificar a intenção de pergunta do indivíduo e direcioná-lo para o diálogo e resposta correta. Contudo, são *chatbots* que utilizam *Machine Learning*, possibilitando que aprendam com experiências de interações anteriores. São também sistemas mais contextuais, podendo estabelecer um diálogo mais personalizado com o usuário. Além disso, tendem a responder perguntas de maneira mais ampla, permitindo, algumas vezes, que o usuário saia um pouco do domínio específico de conhecimento do *chatbot*.

Figura 5: Exemplo fictício de *chatbot* de viagens com Inteligência Artificial



Fonte: O autor (2020)

Nota-se na Figura 5 que o *chatbot* consegue manter um diálogo em linguagem natural no contexto e no domínio no qual ele foi treinado, neste caso, informações sobre viagens. Diferente dos *chatbots* baseados em regras, esses sistemas interagem com o usuário de maneira mais fluida, identificando as intenções de perguntas independentemente da maneira em que a pergunta é feita. Ao perguntar "Quero ir para São Paulo", "Voo para São Paulo" ou "Reservar passagem para SP" o sistema tende a identificar, quando bem treinado, que a intenção do usuário é, provavelmente, realizar uma reserva de viagem para São Paulo. Logo, com base nessa intenção identificada, o usuário é direcionado para um fluxo de diálogo que corresponda a sua necessidade de informação.

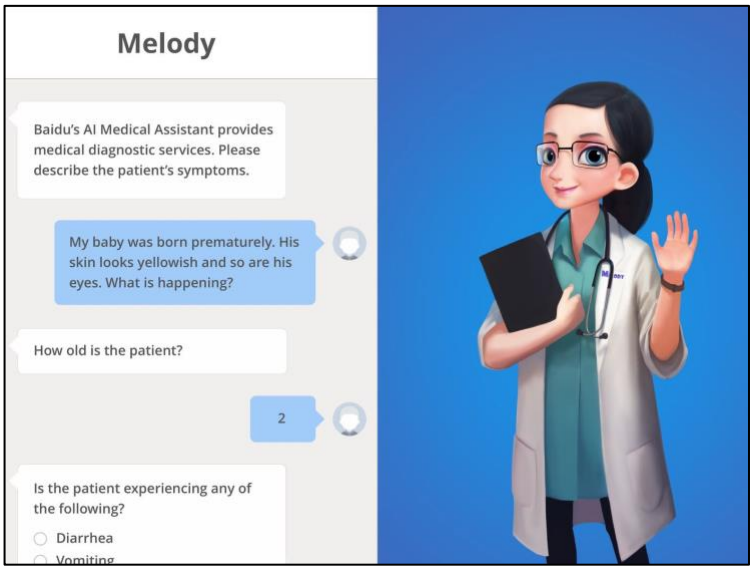
Em conclusão, esses dois tipos de *chatbots* são válidos para a recuperação da informação. Não se pode dizer que um é mais inteligente que o outro, mas sim que cada um tem suas peculiaridades de acordo com o objetivo de uso, além de empregar técnicas distintas no momento do seu desenvolvimento. Os *chatbots* baseados em regras tendem a tornar a comunicação mais direta e objetiva, muitas vezes ajudando a evitar que o usuário tenha que pensar no que perguntar ou como descrever as suas necessidades de informações. Em contrapartida, os *chatbots* baseados em AI permitem melhor compreensão da necessidade do usuário, graças a NPL que é capaz de interpretar diferentes variações de frases em uma mesma solicitação. Existem também os agentes que contemplam ambas características, conhecidos como *chatbots* híbridos (ARTIFICIAL SOLUTIONS, 2020). Em síntese, independente do tipo de *chatbot*, seja ele com ou sem inteligência artificial, todos compartilham de um ou mais elementos fundamentais, tanto para o seu desenvolvimento como para o seu funcionamento; a saber: Persona, Entidade, Intenções, Diálogo e Canal.

Persona

Como o próprio nome diz, persona trata da personalidade do *chatbot* (TAKE, 2017). Aqui estão reunidas diversas características da personalidade do sistema, tais como, o tipo de linguagem utilizada; se terá uma abordagem mais formal ou descontraída; o tipo de assunto no qual o *chatbot* será especialista; se é um agente com a personalidade jovem ou não. É na persona que é desenvolvido o *avatar* e definido o nome que será escolhido para o *chatbot*. Nos exemplos citados acima sobre os diálogos, o nome dado para o especialista em viagens foi "Mochileiro" - pessoa que gosta de viajar ao redor do mundo utilizando, normalmente, apenas uma mochila

nas costas. Fica evidente que, para construir uma persona, se faz necessário conhecer, com bastante profundidade, o usuário que irá utilizar o sistema e, além disso, é necessário que o *chatbot* esteja em conexão com as características da marca.

Figura 6: Exemplo de persona (avatar) em um *chatbot* sobre saúde



Fonte: The Verge (2016)

Intenção

De acordo com a documentação do serviço de desenvolvimento de *chatbots* da IBM (IBM CLOUD, 2020), intenções "[...] são objetivos ou metas que são expressos na entrada de um cliente, como responder a uma pergunta ou processar um pagamento de fatura". Intenção, como o próprio nome sugere, é necessidade de informação do usuário identificada em uma pergunta/sentença. É a ideia principal de uma frase enviada pelo usuário em um *chatbot*. É por meio dessas intenções que esses agentes, em especial aqueles com Inteligência Artificial, conseguem identificar a necessidade do usuário para que a conversa seja direcionada para o fluxo de diálogo adequado e, com isso, a ação desejada pelo usuário seja executada. Normalmente, a intenção vem em forma de um verbo ou ação expressado ou não em uma frase. Exemplos: "Quero saber a temperatura de São Paulo"; "Quantos graus está fazendo hoje?", "Será que amanhã vai chover?", "Domingo é dia de sol ou chuva?". Veja que em ambas frases, a intenção identificada é a mesma: Saber a temperatura.

Ilustração 8: Exemplo de intenção em uma frase

Fonte: O autor (2020)

Entidades

Também de acordo com a IBM (IBM CLOUD, 2020), "as entidades representam informações na entrada do usuário que são relevantes para a finalidade do usuário". Em outras palavras, entidade é sobre o que a intenção se refere, são os elementos mais importantes para filtrar a ação que o usuário deseja executar no *chatbot*. As entidades complementam o sentido das intenções. Se intenções representam verbos (tipo de ação que o usuário deseja), as entidades, basicamente, representam os substantivos. Logo, a pergunta "Vai chover em São Paulo?", tem como entidade o trecho "São Paulo", pois é a palavra chave que contextualiza a intenção (saber a temperatura). Com essa informação, o *chatbot* dará uma resposta mais direcionada sobre as condições climáticas em São Paulo.

Ilustração 9: Exemplo de entidade em uma frase.

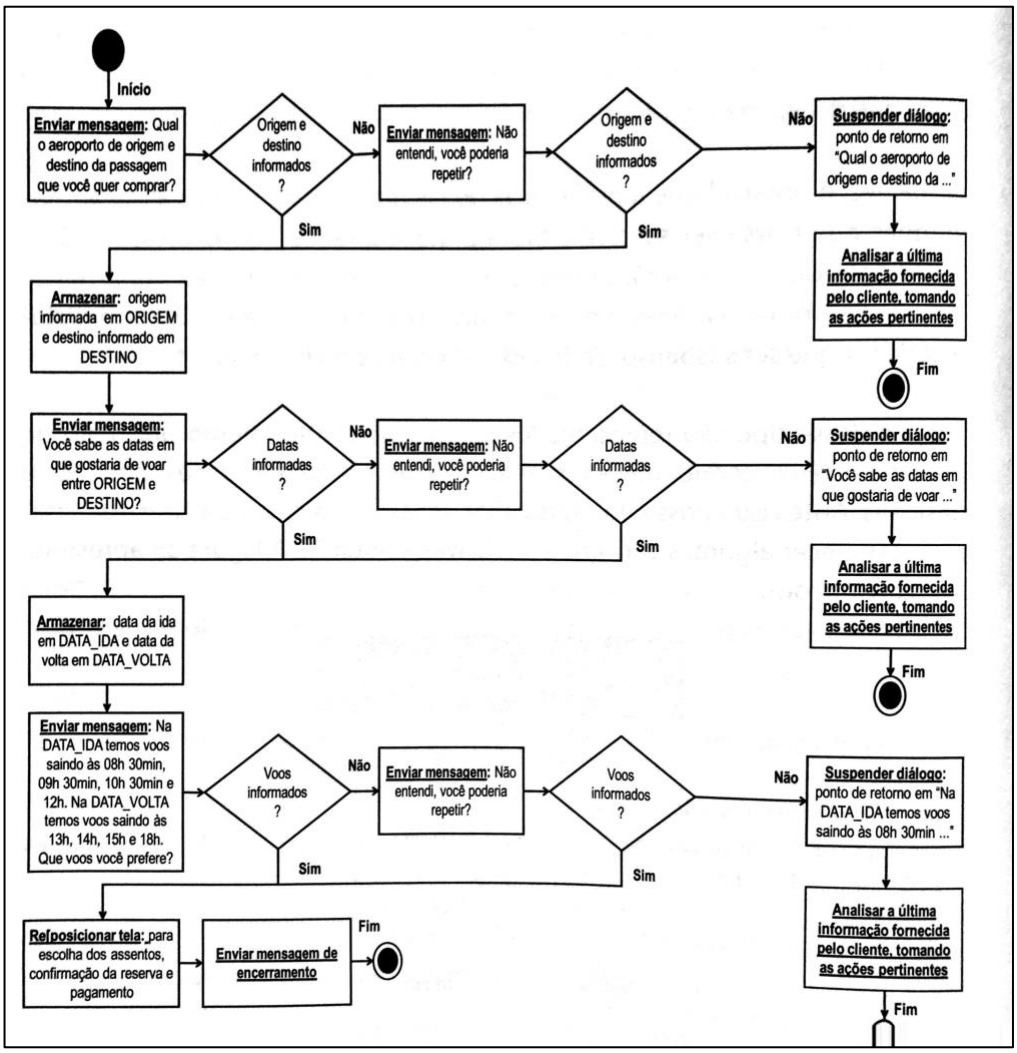
Fonte: O autor (2020)

Diálogo

O diálogo é um elemento fundamental para que a interação com o *chatbot* aconteça. A base de conhecimento, onde é desenvolvida toda a estrutura de respostas de acordo com a questão apresentada pelo usuário (ALENCAR, SCHMITZ E CRUZ, 2018). O diálogo utiliza a intenção do usuário, identificada em sua questão, para que a interação e respostas corretas ocorram. Por exemplo, se o indivíduo pergunta algo como "Quero comer", o sistema provavelmente irá

identificar a intenção, que, neste caso, é o desejo de comer algo. Entretanto, falta um direcionamento, por exemplo, o alimento que o indivíduo deseja comer. Existe a necessidade de se detalhar essa intenção que, neste caso, nada mais é que a entidade. Uma das funções do diálogo é identificar a falta de um complemento na pergunta do sujeito. Dessa forma, ao identificar essa lacuna, o agente poderia perguntar de volta algo como: “O que você deseja comer?” ou “O que você gosta de comer?”. Com a resposta a essas perguntas, o usuário seria direcionado ao fluxo de diálogo que abrangeria a informação ou ação desejada.

Figura 7: Exemplo de uma estrutura de diálogo.



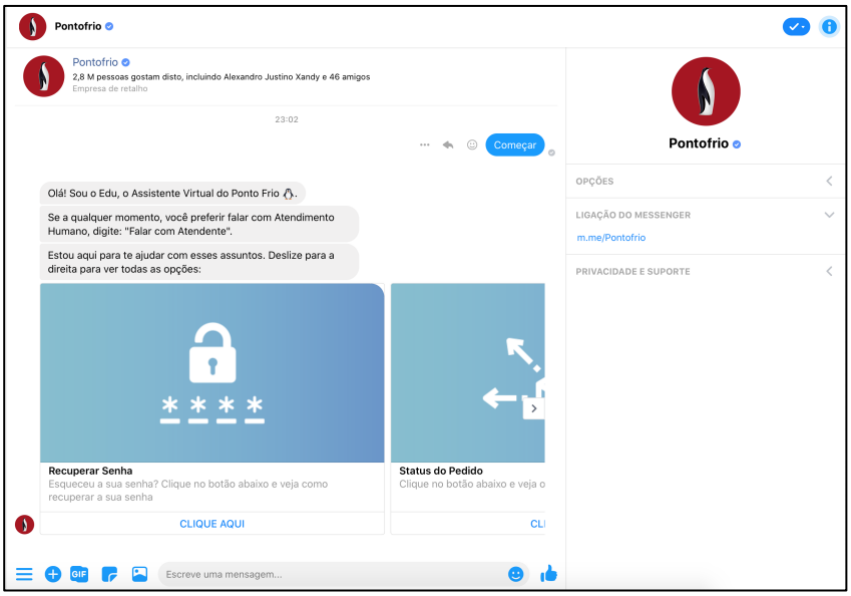
Fonte: Alencar, Schmitz e Cruz (2018)

Canal ou Interface

Por fim, o canal ou interface é onde se estabelece a interação com o *chatbot*. Os exemplos de canais são: WhatsApp, Facebook Messenger, Telegram, Slack, site da empresa. (MICROSOFT,

2018). Normalmente, um agente é desenvolvido em uma única plataforma de desenvolvimento e, após o seu término (definição de intenções, entidade, diálogo), essa plataforma de desenvolvimento permite que o desenvolvedor escolha o canal para hospedá-lo (processo conhecido como *deployment*), seja colocando-o direto em um site ou utilizando um serviço de IM. A Alexa, assistente virtual da Amazon, também pode ser considerada um canal. A plataforma da Amazon (2020) permite que desenvolvedores criem seus próprios *chatbots* (ou *skills*) para serem utilizados por meio da Alexa. Assim, toda a interação é estabelecida por meio de voz.

Figura 8: *Chatbot* do Ponto Frio utilizando o Facebook Messenger como canal



Fonte: Print do Facebook Messenger – Ponto Frio

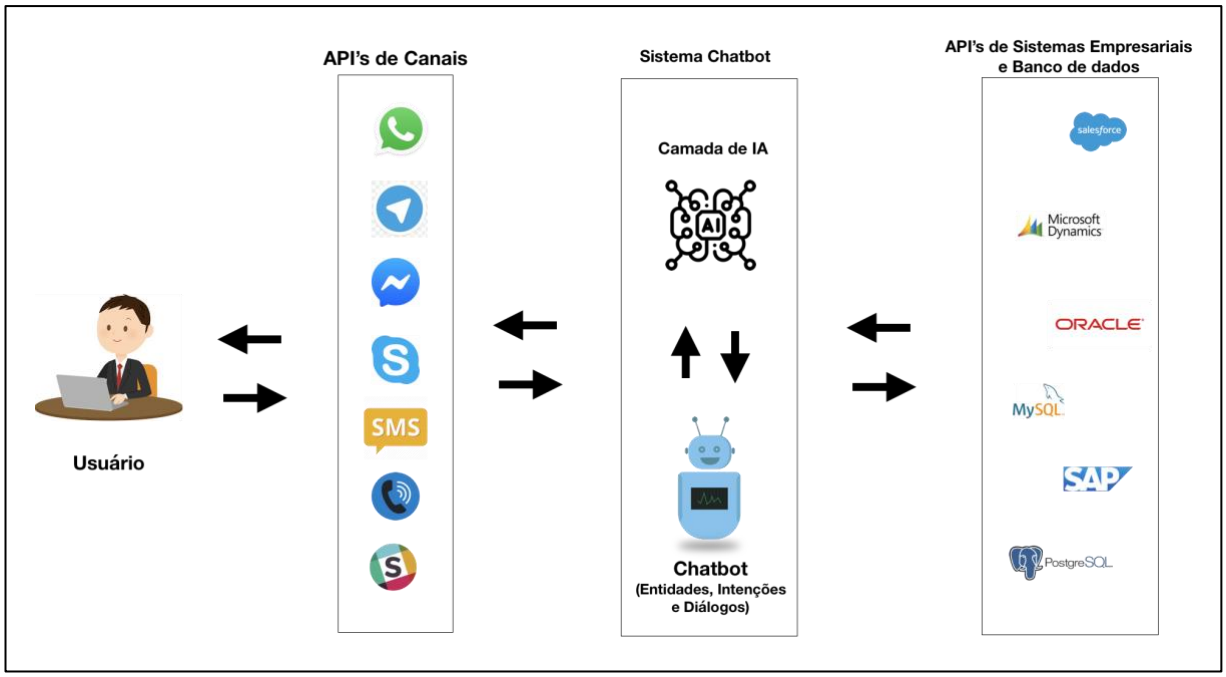
Em suma, esses são alguns elementos fundamentais dentro de um *chatbot*, seja ele simples ou complexo. É claro que, dependendo da plataforma de desenvolvimento, esses elementos aparecerão com nomenclaturas diferentes dos versados aqui. Porém, o mais importante é que eles servirão como base no processo de construção desses agentes. E, nesse sentido, convém entender como é a arquitetura de um *chatbot* sob o viés de suas conexões, isto é, suas integrações com a base de conhecimento, serviços externos, serviços de Inteligência Artificial etc. Essas integrações são comumente realizadas por meio de uma API (*Application Programming Interface*)

De acordo com Michael Meng (2018) API é,

> As interfaces de programação de aplicações (APIs) expõem serviços ou dados fornecidos por um aplicativo de software por meio de um conjunto de recursos predefinidos, como métodos, objetos ou URIs. Ao usar esses recursos, outros aplicativos podem acessar os dados ou serviços sem precisar implementar os objetos e procedimentos subjacentes. As APIs são centrais para muitas arquiteturas de software modernas, pois fornecem abstrações de alto nível que facilitam as tarefas de programação, suportam o design de aplicativos de software distribuídos e modulares e a reutilização de código.

Uma API é um meio que dois serviços de tecnologia utilizam para realizar a comunicação e compartilhamento de suas funcionalidades entre eles. É um processo realizado entre *software-to-software*, ou seja, todo o arcabouço de comunicação entre esses serviços se realiza de forma escondida do usuário. Um exemplo de integração via API é quando se utiliza o mapa do Google dentro de aplicativos que, não necessariamente, foi desenvolvido pelo Google. Por meio da API do serviço de Mapas disponibilizado pelo Google, consegue-se integrar essa funcionalidade dentro de um aplicativo de terceiro. Em outras palavras, o desenvolvedor não precisa criar um serviço de mapa do zero. Esse mesmo raciocínio é utilizado em *chatbots* no momento em que ele é desenvolvido e integrado em uma plataforma de mensagens – como o WhatsApp - ou um sistema empresarial da empresa.

Ilustração 10: Exemplos de integrações em um sistema de *chatbot*



Fonte: O autor (2020)

Na Ilustração 10, existe o usuário que imputa a pergunta no *chatbot*, utilizando um canal conectado por meio de uma API. Essa questão chega ao sistema de *chatbot* e é processada pela

camada de inteligência artificial, seja para o processamento de linguagem natural, a identificação de uma imagem ou a conversão de texto para voz e vice e versa. Dependendo da solicitação do usuário, essa informação pode ser recuperada utilizando a própria base de conhecimento do agente ou um serviço de Inteligência Artificial de terceiros, como o IBM Watson Discovery que realiza a busca em documentos e páginas web quando uma resposta não está pré-definida na estrutura de diálogo do agente. Além disso, essas APIs servem para autenticar o usuário nos sistemas empresariais, como nos casos em que o *chatbot* tenha que fornecer o saldo bancário à um cliente. Não obstante, as APIs podem ser utilizadas para registrar informações em sistemas externos, como em um sistema que registra os pedidos de pizza. Em resumo, a API é uma técnica importante no desenvolvimento de um *chatbot*.

3.4. Desenvolvimento dos *chatbots*

Existem duas maneiras para o desenvolvimento de um *chatbot*. A primeira diz respeito a utilizar uma linguagem de programação e, a partir dela, desenvolver o agente, linha por linha, o que, obviamente, exige um certo conhecimento técnico. A segunda, refere-se a escolher uma plataforma de desenvolvimento de *chatbot* e, por meio de opções intuitivas – como arrastar e soltar -, o desenvolvedor logra construir um sistema do início ao fim. Essa última, em muitas ocasiões, não exige que o indivíduo tenha conhecimento em programação.

Quando se pretende desenvolver um *chatbot* do zero, por meio de codificação, exige-se que o desenvolvedor tenha um conhecimento específico da linguagem de programação em questão. Hoje, existem várias bibliotecas e linguagens que permitem o desenvolvimento de um *chatbot*: AIML, Phyton, PHP, Java, entre outras. Dentre essas opções, uma das mais simples utilizadas para a criação de um agente é a AIML (GALVAO, 2004).

Galvão (2004) explica que:

> A linguagem AIML, baseada em XML, é descrita através de identificadores denominados *tags*. Ela foi desenvolvida tendo como inspiração o modelo de casamento de padrões proposto por ELIZA.

Galvão continua (2004):

> Em AIML, cada regra (categoria) é representada por uma *tag* *<category>*. Essencialmente, cada categoria é constituída por um *componente de decomposição* da

sentença digitada pelo usuário e por um *componente de reconstrução* (ou construção) da sentença devolvida pelo *chatterbot* ao usuário, sendo denominados respectivamente como *pattern* e *template*.

Trata-se de uma linguagem de marcação (estruturada por *tags*) que permite a elaboração de um diálogo de *chatbot*. A estrutura dessa linguagem inicia-se com a tag <*category*> que representa o início de um diálogo. Em seguida vem a tag <*pattern*> que simboliza a mensagem inserida pelo usuário. Finalmente, a tag <*template*>, que constitui a resposta do usuário à aquela pergunta/informação inserida por ele.

Ilustração 11: Ilustração de diálogo simples com a linguagem AIML

```
<category>
<pattern>Bom dia!</pattern>
<template>Bom dia, tudo bem com você?</template>
</category>

<category>
<pattern>Quero viajar</pattern>
<template>Para você deseja viajar? São Paulo / Rio de Janeiro</template>
</category>
```

Fonte: O autor (2020)

Na ilustração 11, existem dois diálogos simples, cada um representado entre as *tags* <*category*> e </*category*>. Ao executar esse código, o *chatbot* funcionaria da seguinte maneira. A partir do momento que alguma informação fosse digitada pelo usuário, o código seria "varrido" até comparar a informação entre as tags <*pattern*> e </*pattern*> que fosse igual àquela digitada pelo usuário. Dessa forma, ao digitar "**Bom dia!**", o código processaria e encontraria a informação correspondente dentro da estrutura <*pattern*>**Bom dia**!</*pattern*> e, logo em seguida, retornaria para o indivíduo a resposta entre as *tags* <*template*> e </*template*> que, neste caso, seria "**Bom dia, tudo bem com você?"** O mesmo racional seria utilizado para quando o usuário digitasse "**Quero viajar**".

Em síntese, esses são apenas alguns exemplos de *tags* que podem ser utilizadas dentro da linguagem AIML. Existem diversas outras que auxiliam na acurácia da recuperação da informação. Como se pode observar, desenvolver um *chatbot* exclusivamente via linguagem de programação é um processo que pode exigir bastante trabalho, já que cada diálogo representa diversas linhas de comando. Além disso, no caso da AIML, trata-se de uma linguagem que não tem um processamento de linguagem natural embutido, ou seja, todo o diálogo é estruturado entre *match* com a palavra-chave digitada pelo usuário. Fazendo necessária a criação de várias

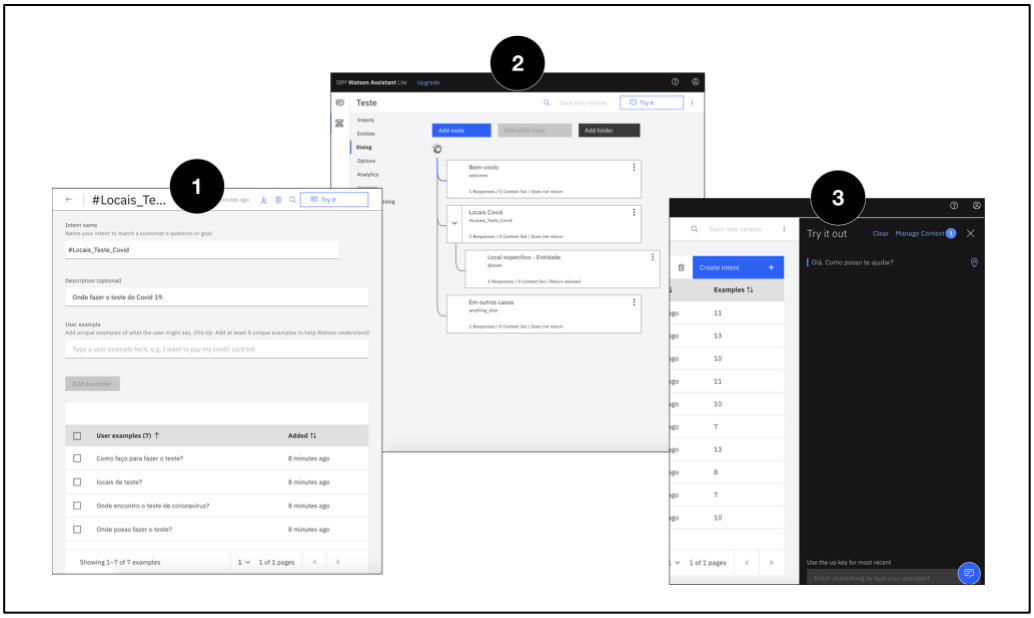
combinações para uma simples mensagem de saudação, já que o indivíduo poderá faze-la de diversas maneiras. Para se ter uma ideia, a base de conhecimento do *chatbot* A.L.I.C.E., que foi desenvolvido com a linguagem AIML, utilizou cerca de 42 mil *tags* <*category*>, ou 42 mil conjuntos de perguntas e respostas (HENN, 2006).

Outra forma de desenvolver um *chatbot* é por meio de plataformas de desenvolvimento. O acesso a essas plataformas ocorre com o uso de serviços em computação em nuvem, conhecidos como SaaS (*Software as a Service*) ou PaaS (*Platform as a Service*). Toda a programação e desenvolvimento de um agente acontece através de interações, como arrastar, digitar, inserir. Isto é, o *chatbot* continua sendo desenvolvido por meio de códigos de programação, bem como qualquer outro *software*. Entretanto, essas plataformas são utilizadas como um intermediador para facilitar o desenvolvimento do agente para aquele profissional que não tem domínio em programação. Além disso, essas plataformas normalmente vêm com camadas de inteligência artificial e opções de conexões nativas com outros sistemas e canais (utilizando-se de APIs). Para exemplificar, três plataformas serão analisadas: IBM Watson Assistant, DialogFlow e Chatfuel.

IBM Watson Assistant

O IBM Watson Assistant é um dos diversos serviços de Inteligência Artificial da IBM disponibilizado por meio de sua computação em nuvem, o IBM Cloud (2020). O IBM Watson Assistant, é um serviço em nuvem (não exige nenhuma instalação de programa na máquina do usuário) que permite a criação de um *chatbot* com Inteligência Artificial em questão de minutos. Todo o manuseio da plataforma é realizado diretamente pelo browser de navegação do indivíduo, seja Google Chrome, Safari, Internet Explorer, Firefox. Além disso, é uma plataforma que possui versões gratuitas e pagas, possibilitando que qualquer pessoa construa um *chatbot* com características simples chegando a obter soluções mais robustas.

Figura 9: Telas de exemplos do IBM Watson Assistant



Fonte: Adaptado de IBM Cloud (2020)

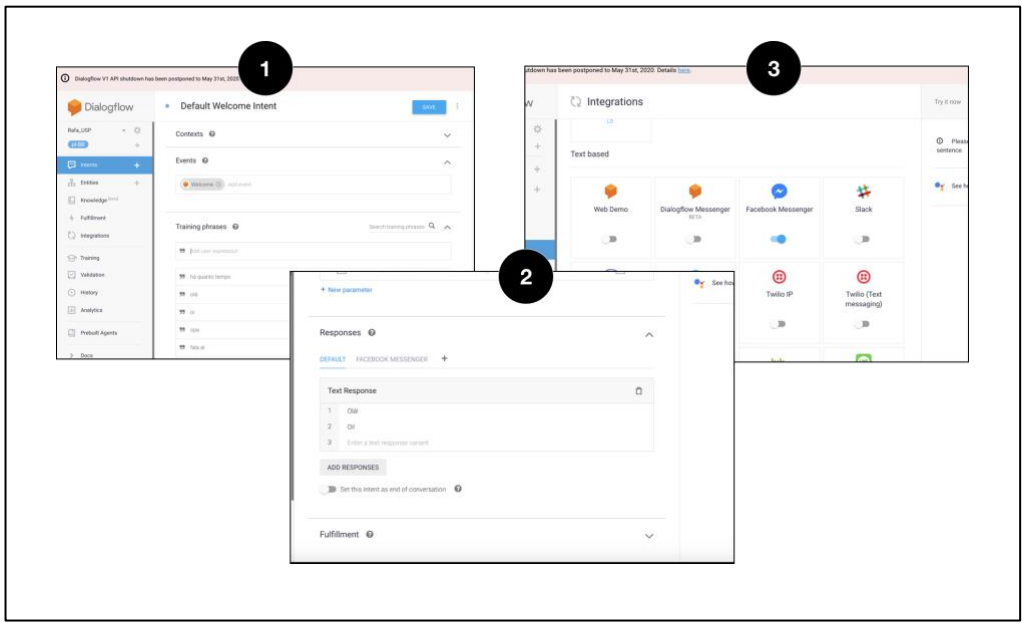
Na plataforma, o desenvolvedor poderá criar as intenções para o *chatbot* e treiná-las por meio de frases exemplos (Figura 9 – referência 1). Por meio de *Machine Learning*, o sistema é capaz de entender que frases como “Bom dia”, “Boa tarde”, “Oi, tudo bem?” “E aí, blz?” ou “Opa, como vai?” significam uma só intenção; que neste caso, é fazer uma saudação. Com base nestas intenções identificadas, o desenvolvedor passa a criar diálogos que, como visto, é a base de conhecimento de um *chatbot*. Essa base é construída em uma estrutura de árvore de decisão, que leva o usuário a informação desejada (Figura 9 – referência 2). No entanto, o diálogo poderá se tornar ainda mais detalhado com o uso de entidades. Em suma, a plataforma também permite que o *chatbot* seja testado em tempo real pelo desenvolvedor (Figura 9 – referência 3). Tornando possível verificar se as intenções e as entidades estão sendo reconhecidas e as perguntas correspondidas.

DialogFlow

O DialogFlow é também um serviço de *chatbot* baseado em nuvem, isto é, não há necessidade de instalação de software na máquina do usuário e, além disso, toda a operação é realizada por intermédio de um browser de navegação. É um serviço oferecido pelo Google Cloud (2020) que permite que o profissional desenvolva um *chatbot* em questão de minutos. Trata-se de uma plataforma que tem uma estrutura bastante parecida com o IBM Watson Assistant, onde conta

com as entidades, intenções e regras de diálogos; bem como um ambiente para teste em tempo real.

Figura 10: Telas de exemplos do DialogFlow



Fonte: Adaptado de Google Cloud (2020)

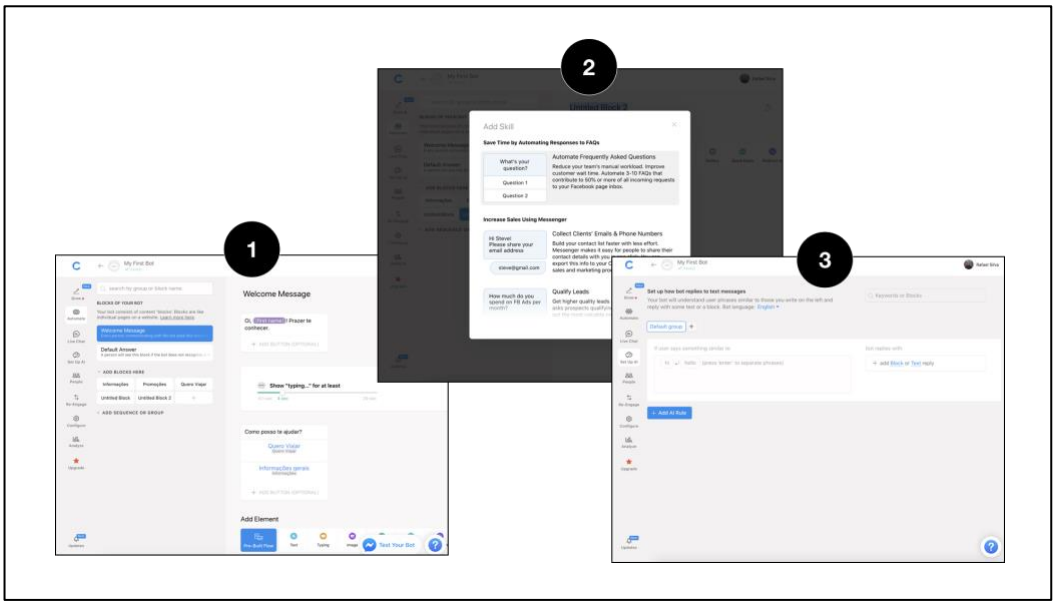
Como o IBM Watson Assistant, o DialogFlow permite desenvolver um *chatbot* com inteligência artificial. Desta forma, as intenções detectadas na frase de um usuário podem ser treinadas com frases exemplos (Figura 10 – Referência 1) para que todas elas sejam identificadas como uma única intenção, facilitando o desenvolvimento do diálogo (Figura 10 – Referência 2). Além disso, o DialogFlow possibilita a integração em diversos canais de mensagens instantâneas, tais como Facebook Messenger, Slack, Skype. Contudo, é possível ter uma integração direta com sistemas de telefonia, o que permitiria uma interação com o *chatbot* por meio de uma URA de atendimento (Figura 10 – Referência 3). Em síntese, uma vez finalizado o desenvolvimento, bastaria que o profissional fizesse o *deployment* no canal desejado. Outro ponto interessante dessa plataforma é a API nativa de *Speech Recognition* do Google Assistant. Com isso, o desenvolvedor poderia criar um *chatbot* onde a interação aconteceria por voz, seja utilizando a técnica *Speech-to-Text*, como *Text-to-Speech*.

Chatfuel

O Chatfuel (2020) é uma plataforma focada no desenvolvimento de *chatbots* para o Facebook Messenger. Trata-se de um serviço onde é possível criar agentes de forma gratuita e bastante

simples. Algumas empresas que utilizam o Chatfuel são: NBA, ABC News, T-Mobile, Adidas. O processo de desenvolvimento pode ser realizado em questão de minutos, além de não exigir nenhum tipo de experiências prévias ou conhecimento em linguagem de programação por parte do desenvolvedor. É, portanto, uma plataforma bastante intuitiva do tipo baseado em regras e recomendada para pessoas que estão começando a desenvolver um agente conversacional.

Figura 11: Algumas telas do ChatFuel



Fonte: Adaptado de Chatfuel (2020)

Comparado com o IBM Watson Assistant e DialogFlow, o Chatfuel possui uma estrutura distinta para a construção de *chatbots*. Todo o diálogo é organizado por blocos (Figura 11 – Referência 1). Dessa maneira, como as intenções, cada bloco se concentra em um assunto específico dentro do domínio do agente. Dentro de cada bloco, o desenvolvedor consegue inserir o tipo de interação com o usuário, seja por meio de botões, links ou imagens. Os botões, por exemplo, podem levar o usuário para um bloco (assunto) correspondente, bastando que o indivíduo clique sobre eles. As interações também oferecem algumas opções de blocos ou *skills* pré-definidas (Figura 11 – Referência 2) como um fluxo de diálogo exclusivo para a captação de dados de clientes para alguma ação de marketing. Em síntese, apesar de ser uma plataforma para a construção de *chatbots* do tipo baseado em regras, ela também oferece a opção para que o profissional utilize uma camada de inteligência artificial (especificamente o processamento de linguagem natural). Semelhante ao IBM Watson Assistant e DialogFlow, é necessário que o

sistema seja treinado para que o *chatbot* direcione o usuário para o bloco correto (Figura 11 – Referência 3).

Como visto, existem diversas maneiras de desenvolver um *chatbot*, seja utilizando uma linguagem de programação ou plataformas de desenvolvimento em nuvem. Essas plataformas possibilitam que o profissional possa desenvolver seus agentes em questão de minutos e com pouco conhecimento em linguagem de programação. Além disso, um único desenvolvimento permite o *deployment* em diversos canais (Facebook, WhatsApp, Slack, Telegram). Isto é, o desenvolvedor não precisa criar um *chatbot* diferente para cada canal. Outro ponto interessante dessas plataformas é a possibilidade de ter acesso a um painel que mostra as interações e dados estatísticos de utilização do agente. É possível monitorar se as respostas de fato atenderam a solicitação do usuário e quais foram as perguntas não entendidas pela ferramenta. Com base nesses dados, esses agentes podem ser melhorados para que as próximas perguntas possam ser respondidas de forma acurada.

Em resumo, são plataformas onde qualquer pessoa poderá desenvolver um *chatbot*, mesmo aquelas que não tenham domínio em programação. É importante ressaltar que plataformas como o IBM Watson Assistant e DialogFlow, também permitem que desenvolvedores, com conhecimentos avançados em linguagem de programação, consigam manipular o código fonte dessas aplicações. O *chatbot* poderá se tornar ainda mais personalizado, não precisando limitar-se as funções de desenvolvimento oferecidas pela plataforma. É por essas e outras que, pela sua facilidade de desenvolvimento e uso, os *chatbots* estão sendo vistos como uma importante tendência em diversos contextos de recuperação da informação, como será visto a seguir.

3.5. Popularidade e principais tendências

Conforme versado no começo deste capítulo, os *chatbots* surgiram desde a década de 1960, porém somente nos últimos anos essa tecnologia vem se potencializando (ALENCAR, SCHMITZ E CRUZ, 2018). Muitos são os motivos para a sua popularidade. Segundo Dale (2016), uma das razões foi o crescimento de aplicativos de conversas, como WhatsApp e Telegram. Segundo o autor, a popularização e o surgimento desses aplicativos de mensagens instantâneas contribuíram para que adaptação e o uso dos *chatbots* se tornasse algo natural entre as pessoas. Em outras palavras, o uso e a interação (seja por mensagens de texto ou voz) com

esses aplicativos se tornou algo natural no dia a dia da população. Essa naturalidade traz um ambiente quase perfeito para os *chatbots*, já que o usuário não precisaria, necessariamente, aprender a interagir com esses agentes. Basta conversar com eles, do mesmo modo que é feito em outros aplicativos. Não obstante, não se faz necessário baixar um aplicativo específico para o seu uso, já que muitos desses agentes já são desenvolvidos para serem utilizados nos aplicativos que os usuários já utilizam.

Já Alencar, Schmitz e Cruz (2018) argumentam que outros fatores que contribuíram para a popularização dos *chatbots* foram os avanços e melhorias em tecnologias como a Inteligência Artificial, sobretudo o Processamento de Linguagem Natural. Como abordado anteriormente, o Processamento de Linguagem Natural permite que os *chatbots* reconheçam de forma natural a linguagem falada pelo ser humano. Esses algoritmos tendem a entender e satisfazer de forma mais acurada as necessidades de informação do usuário. Apesar de, a melhoria e o barateamento do poder de processamento das máquinas corroborarem para o surgimento de tecnologias como PLN, já que seus algoritmos exigem um desempenho rápido e versátil que, somente agora, surgem de forma escalável nos computadores atuais.

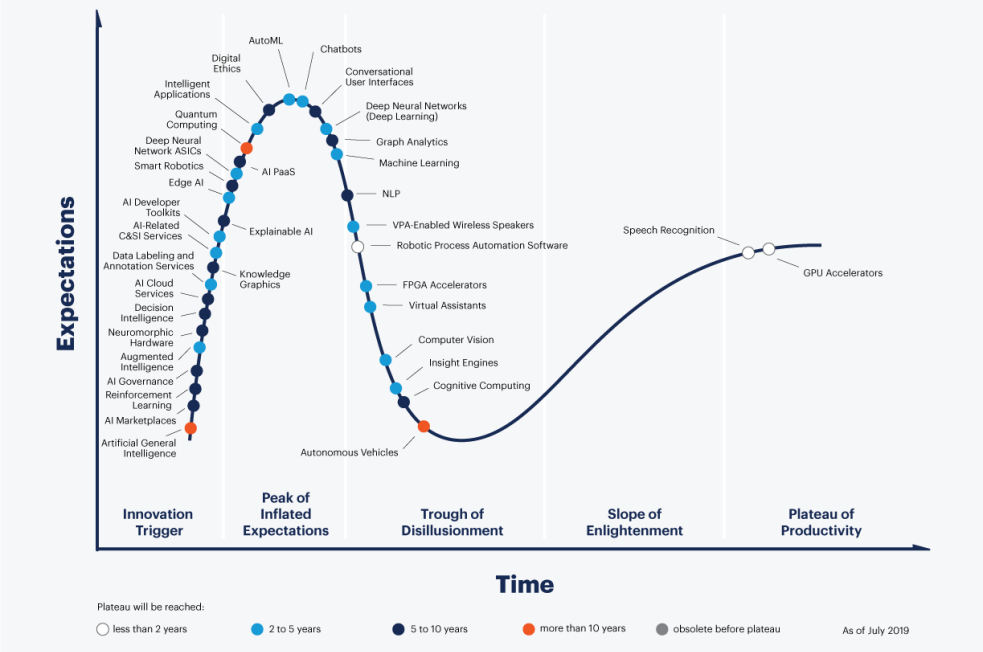
Analisando o momento atual da sociedade, é possível inferir que, outro fator que contribuiu para a popularização dos *chatbots* é a necessidade dos usuários em obter informações organizadas em tempo real, em um momento onde diversas delas surgem a cada momento (ALENCAR, SCHMITZ E CRUZ, 2018). Esse cenário fica mais evidente quando se analisa os serviços de atendimento ao cliente, pois o perfil de consumo das pessoas mudou (KOTLER, 2017). As instituições precisam oferecer contato e acesso rápido a informação, tornando-se algo obrigatório para a sua sobrevivência. Isto é, garantir um atendimento e acesso rápido a informação é algo essencial, pois, caso contrário, a insatisfação fará com que esse indivíduo procure outra empresa ou serviço. Nesse sentido, os *chatbots* surgem como uma solução a essa demanda, já que se pode oferecer um atendimento 24h, permitindo que o usuário esclareça suas dúvidas em tempo real, sem, necessariamente, depender de um atendimento humano, limitado por um horário comercial. Por esses motivos que os *chatbots* estão se tornando uma tendência.

Segundo a Gartner (2019), até 2021, 1 a cada 6 serviços de atendimento ao consumidor irá utilizar algum serviço de inteligência artificial (como os *chatbots*) para atender os seus clientes. Já a Oracle (2016) prevê que ainda em 2020, 80% das empresas planejam utilizar *chatbots* em suas estratégias de negócios. Outra pesquisa da Juniper Research (2017) estima que até 2022

os bancos irão automatizar os seus processos em até 90% com o uso de *chatbots*. Em relação ao valor monetário desses agentes, de acordo com a Markets and Markets (2019), esse mercado pretende crescer mais de U$ 15 Bilhões de dólares até 2024. Em resumo, esses dados corroboram para o que apresenta a Gartner em seu *Hyper Cycle*.

No decorrer de alguns anos de pesquisa, a Gartner (2020) identificou um padrão entre o surgimento e amadurecimento de uma tecnologia e, com base nisso, criou o *Hype Cycle*. O *Hype Cycle* é uma das principais metodologias de pesquisas de tendências de tecnologias desenvolvida pela Gartner. O gráfico fornece uma representação visual da maturidade e adoção de algumas tecnologias e aplicações, mostrando como elas tendem a evoluir ao longo do tempo, desde sua concepção até a sua maturidade. É, portanto, uma pesquisa muito utilizada pelas empresas, já que as ajudam na tomada de decisões sobre a adoção e investimento de determinada tecnologia.

Figura 12: Gartner Hype Cycle para Inteligência Artificial



Fonte: Gartner (2019)

O gráfico é divido em 5 fases: *Innovation Trigger*, categoriza as tecnologias que começam a ganhar visibilidade da mídia, porém ainda não estão prontas para uso comercial; *Peak of Inflated Expectations* mostra o auge da popularidade de uma tecnologia e as expectativas geradas ao redor dela; *Trough of disillusionment* é a fase que indica certo exagero em relação a tecnologia e a mesma passa a ser esquecida, caso novas implementações não sejam viabilizadas;

Slope of enlightenment indica que, após a fase de glórias e declínio, a tecnologia passa a ganhar maturidade, melhorias e formas para melhor aplicação e uso; *Pleateu of productivy* é a fase final e indica a madurez da tecnologia e sua adesão de forma ampla por empresas e pessoas.

> Embora já estejam em uso, os *chatbots* ainda têm um enorme potencial para impactar o número de agentes de serviço empregados por uma empresa e como o atendimento ao cliente é realizado. À medida que os *chatbots* começarem a "aprender" os desejos dos usuários, os casos de uso da tecnologia serão expandidos para ajudar na integração, treinamento, produtividade e eficiência dentro do local de trabalho. (GARTNER, 2019)

Analisando o gráfico, o *chatbot* encontra-se em seu auge (*Peak of Inflated Expectations)*. A Gartner estima que esses agentes conversacionais estarão nessa fase entre 2 a 5 anos, período esse que os *chatbots* terão que provar a sua real utilidade, antes de entrar em sua fase de queda – caso isso aconteça. Portanto, esse é o momento do qual as pessoas e as empresas estão criando, testando e comprovando a eficiência desses sistemas. É o momento que esses agentes serão utilizados em diversos contextos, em especial aqueles de recuperação de informação.

> O futuro nunca esteve tão maduro para o sucesso do *chatbot*. Os canais de mensagens estão adotando *chatbots* e fornecendo a eles recursos técnicos avançados. As marcas estão cada vez mais abertas às vantagens que essa tecnologia os ajuda a obter. E há uma dinâmica na situação atual que pode muito bem levar a um grande salto no futuro para os *chatbots*, além de tudo o que os aplicativos alcançaram. As plataformas de desenvolvimento do *chatbot*, como o SnatchBot e o Chatfuel, possibilitam a criação de um *chatbot* por qualquer pessoa. Estamos nadando nas águas subindo de um *tsuami* de criação de *chatbot* e, se milhões de *chatbots* de fontes coligadas forem vinculados de uma maneira que lhes permita aprender com suas interações ... bem, isso seria a base de uma revolução mais profunda do que a internet. (MULDOWNEY, 2017 – tradução o autor)

Como apresentado, a história dos *chatbots* mostra que apesar de serem bastante populares hoje, a sua idealização existe há muito tempo. Foi visto também que o uso da inteligência artificial os torna ainda mais eficientes no processo de recuperar a informação e, por conseguinte, tendem a melhorar a satisfação dos usuários ao compreender as suas necessidades de informação. Além disso, esses *chatbots* podem ser incorporados e potencializados com o uso de APIs, permitindo uma vasta interoperabilidade de informações entre sistemas e banco de dados. Por fim, quando o objetivo é construir um chatbot pequeno e de diálogo simples, as diversas ferramentas de desenvolvimento existentes torna esse desenvolvimento simplificado por qualquer pessoa, mesmo aquelas que não tenham nenhum conhecimento de linguagem de programação. Em suma, essas evidências sugerem que hoje a sociedade vive na era dos *chatbots*.

4. BATENDO UM PAPO COM A INFORMAÇÃO: ESTUDOS DE CASOS

A ideia deste capítulo é evidenciar, por meio da descrição de casos, como os *chatbots* podem colaborar com a Ciência da Informação no âmbito da Recuperação da Informação. Alguns casos foram analisados por este autor, baseado na interação real com agentes em seus respectivos contextos. Entretanto, outros agentes foram analisados pautados em documentos, notícias e artigos que os descrevem, já que o acesso a eles foi limitado ou simplesmente o projeto não existia mais.

> [...] se o cliente não tiver certeza de quais termos procurar, ou onde os termos de pesquisa usados são muito comuns, ele poderá ficar frustrado com a perda de tempo navegando no site por conta própria (Muldowney, 2017 – Tradução nossa).

A citação elucida um pouco os problemas de recuperação da informação nos dias atuais. O objetivo é descrever e analisar o uso dos *chatbots* em diferentes contextos, sempre com o viés de entender como uma conversa em linguagem natural pode ser eficiente na busca por uma informação. Eficiente aqui, entende-se pelo fato de a recuperação da informação ser realizada de maneira mais simples e menos burocrática, evitando, por exemplo, que o usuário navegue por diferentes páginas em um site para encontrar a informação que necessita. Ademais, tem por propósito de refletir como a diversidade de canais ajudam na interação com esses agentes.

O método adotado aqui trata-se dos estudos de casos múltiplos. Na literatura, não há um critério formal para a seleção dos casos. Entretanto, os casos a seguir foram escolhidos com o intuito de trazer atualidade, evidências e responder parte da pergunta problema dessa pesquisa, pois como comenta Muller (citar): “A escolha e identificação da unidade de análise [casos] está associada ao problema ou questão da pesquisa”. Assim, escolheu-se um caso em cada indústria com o objetivo de realizar algum grau de generalização e comparação que permitissem evidenciar como os *chatbots* facilitam a recuperação da informação em diversos contextos de uso.

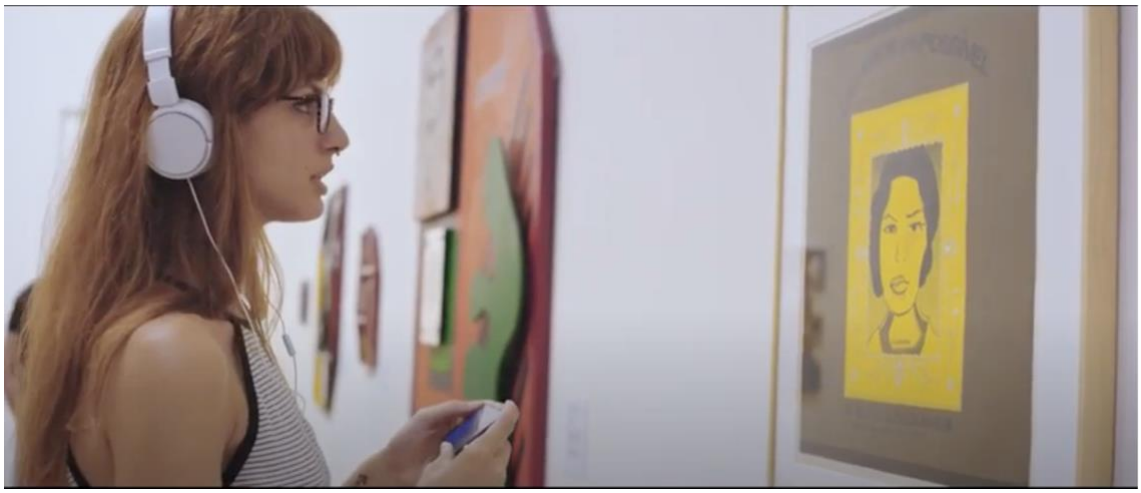
4.1. Quem é o autor da obra? – Estudo de Caso: A voz da Arte

Museus são instituições sem fins lucrativos que têm por objetivo conservar, investigar, comunicar, interpretar e expor, para fins de preservação, estudos, pesquisa, educação,

contemplação e turismo, conjuntos e coleções de valor histórico, artístico, científico, técnico ou de qualquer natureza cultural, aberto ao público e a serviço da sociedade e de seu desenvolvimento (LEI 11.904/09). Acontece que, interpretar e dialogar com obras encontradas nessas instituições, nem sempre ocorre de maneira fácil, tendo como efeito que pessoas deixem de ir ou não tenham interesse em museus, seja por não entenderem as obras encontradas nesse local ou simplesmente por achar o ambiente desinteressante (OI FUTURO, 2019). Nesse contexto, as Tecnologias da Informação e Comunicação, sobretudo a Inteligência Artificial, trazem soluções que, a priori, parecem auxiliar as pessoas a terem melhor interação e entendimento em relação ao que essas expressões artísticas querem comunicar (CHIOVATTO, 2019). Um exemplo de solução, nesse sentido, foi o projeto desenvolvido na Pinacoteca de São Paulo: A Voz da Arte.

O projeto "A Voz da Arte" foi desenvolvido pela IBM em parceria com a agência Ogilvy & Mather e a Pinacoteca – Museu de Arte Contemporânea de São Paulo - no ano de 2017. Tratava-se de um *chatbot* que interagia com os visitantes respondendo perguntas referentes a algumas obras selecionadas no museu (CHIOVATTO, 2019), como a obra Saudade, de Almeida Junior de 1899. O projeto funcionava da seguinte forma: Na chegada à Pinacoteca, o visitante recebia um fone de ouvido e um smartphone com um aplicativo de *chatbot* instalado. Ao andar pelo museu, os visitantes recebiam notificações alertando-os da proximidade de alguma obra selecionada para interação. A partir desse momento, o indivíduo poderia interagir com a obra perguntando para o *chatbot* informações relacionadas a: data da obra; informações sobre o autor; contexto da obra; dúvidas sobre algum elemento no quadro; localização da obra etc. Todas essas perguntas e respostas eram realizadas por meio de interação com texto ou voz.

Figura 13: Momento que mostra a interação com o IBM Watson na Pinacoteca



Fonte: A voz da Arte Youtube (2020)

O *chatbot* foi desenvolvido com o uso do IBM Watson Assistant, que, como visto, trata-se de uma plataforma de desenvolvimento de *chatbots* que permite adicionar camadas de Inteligência Artificial. Dessa forma, ao realizar perguntas para o *chatbot*, essa informação em formato de áudio era convertida em texto (*Speech-to-text*) para que, logo em seguida, passasse por uma camada de Processamento de Linguagem Natural e identificação de intenções contidas nas perguntas (BATH, 2017). Ao identificar a intenção, verificava-se a melhor resposta em uma base de conhecimento com informações pré-definidas. Essa resposta era convertida novamente em áudio (*Text-to-Speech*) e devolvida ao usuário. O sistema foi constantemente treinado por curadores da Pinacoteca (CHIOVATTO, 2019). Esse treinamento incluiu exemplos de perguntas, levantamento das principais intenções e entidades e, principalmente, as respostas à essas perguntas. Não obstante, o *chatbot* era frequentemente realimentado com novas respostas às perguntas que, outrora, não tinham sido respondidas.

Esse projeto proporcionou uma nova maneira de interagir com obras de artes em um museu. Com o uso do Processamento de Linguagem Natural, o público poderia obter informações instantâneas sobre as obras, realizar perguntas em formato de áudio, como se estivessem perguntando à um curador de artes. Além disso, o *chatbot* também permitia que pessoas com deficiência auditiva pudessem interagir com as obras; nesse caso, por meio de perguntas realizadas em formato de texto. De certo, existem diferentes formas para que uma arte dialogue com o público, seja por meio de textos de parede, educadores, folhetos, jogos, áudio-guias. (CHIOVATTO, 2019). Entretanto, fica evidente que os *chatbots* podem ser uma nova maneira

de complementar esse diálogo, pois eles facilitam a recuperação da informação em um ambiente que, no passado, poderia ser intimidador para muitas pessoas com pouco conhecimento ou contexto para interpretar uma obra de arte.

4.2. O que significa esse sinal? – Estudo de Caso: Manual Cognitivo

Manuais de instruções são documentos que auxiliam o usuário no manuseio e uso de um equipamento, objeto, ferramentas, programas de computadores, entre outros. Porém, apesar de muitos manuais virem com imagens e outros tipos de ilustrações, o seu uso e interpretação é, muitas vezes, complicado. Um exemplo disso são os manuais de automóveis. Em uma matéria publicada pelo jornal Folha de São Paulo, pelo jornalista Marco Antônio Rocha (2018), é dito que "cinquenta anos atrás, o manual de bordo de um carro tinha cerca de 50 páginas. Hoje, os guias de veículos feitos no Brasil têm mais de 350 páginas, todas repletas de termos técnicos que mudam de marca para marca. Desta forma, a recuperação de informação nesses documentos se torna prejudicada, sobretudo quando se trata de um usuário leigo ou um indivíduo que simplesmente não tem tempo e paciência para ler uma miríade de páginas para encontrar as informações que precisa. Nesse sentido, o uso de ferramentas tecnológicas, como os *chatbots*, parece ser uma opção para facilitar a busca de informações nesses documentos. Um exemplo disso é o projeto Manual Cognitivo da Volkswagen.

O Manual Cognitivo é um projeto de *chatbot* que, atualmente, auxilia os usuários do carro Virtus, da Volkswagen, a utilizarem melhor o seu veículo (RIBEIRO, 2018). Utilizando o aplicativo "Meu Volkswagen", os usuários podem interagir com um agente, por voz ou texto, realizando perguntas e esclarecendo diversas dúvidas sobre a utilização do veículo, tais como: capacidade; informações técnicas; luzes de alertas no painel. Para os casos onde o usuário tenha dúvidas em relação as luzes no painel, o *chatbot* pede para que o indivíduo tire uma foto desse sinal. Com o auxílio do smartphone, basta que o usuário fotografe o sinal luminoso e, em tempo real, o sistema identifica o sinal e responde ao usuário o que significa a luz acesa no painel do carro.

Figura 14: Funcionamento do *chatbot* – Manual Cognitivo

Fonte: Manual Cognitivo Youtube (2020)

O Manual Cognitivo também foi desenvolvido em parceria com a IBM (RIBEIRO, 2018), logo, utilizou-se o IBM Watson Assistant para a sua criação. Portanto, tal como explicado no capítulo anterior, o sistema permite que o desenvolvedor utilize recursos de inteligência artificial no *chatbot*. Desta forma, uma das técnicas de inteligência artificial utilizada nesse projeto foi o *Image Recognition*. Esse sistema permite que o usuário tire uma foto para, logo em seguida, essa foto ser identificada e vinculada a uma intenção de pergunta. Com isso, ao identificar essa intenção, tal como seria em um texto, o agente retorna ao usuário uma resposta que mais se aproxima a informação da foto identificada por meio de inteligência artificial (IBM, 2020). É importante destacar o trabalho de curadoria de conteúdo e organização de fluxos de mensagens nesse *chatbot*, pois todas as respostas dadas ao usuário contêm riqueza de detalhes que o auxiliam a satisfazer as suas necessidades de informação. Esses detalhes incluem respostas no formato de vídeo, imagens, links, entre outros (IBM, 2020).

É evidente que o índice, uma das técnicas de organização e recuperação da informação, auxilia o usuário a encontrar uma informação desejada, entretanto, trata-se de uma tarefa não trivial em algumas situações. Por exemplo, se o usuário deseja buscar por uma informação indecifrável contida em uma imagem, como um sinal luminoso desconhecido em um painel de um carro, o índice, certamente, não será suficiente para recuperar a informação que satisfaça a sua dúvida. Logo, nesse contexto, os *chatbots* parecem ser uma ótima opção de recuperar esse tipo de informação, já que bastaria que o indivíduo tirasse uma foto para que uma intenção fosse

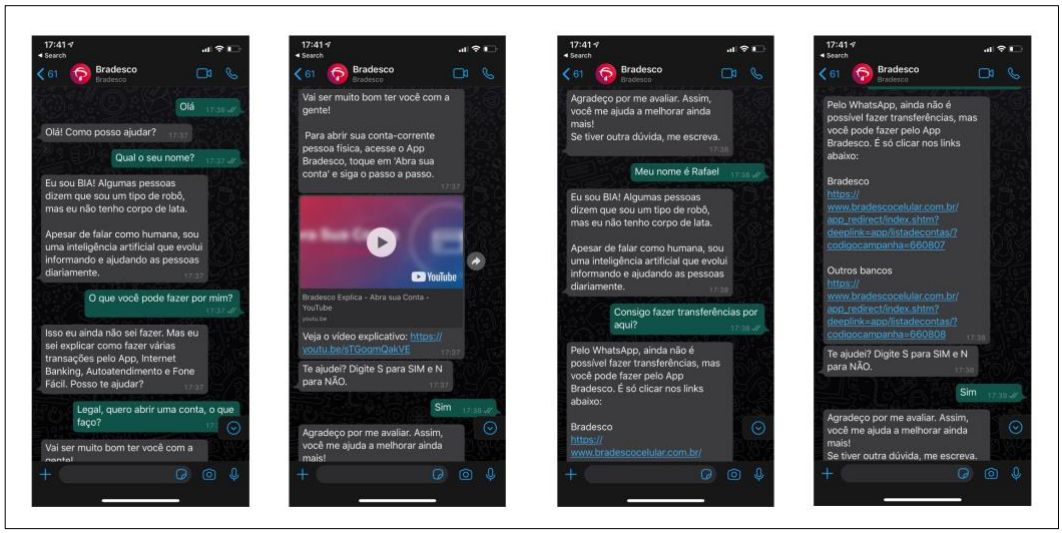
identificada a partir da imagem e, logo em seguida, uma resposta fosse retornada a esse sujeito. Da mesma forma, ao utilizar interação por voz em linguagem natural, a recuperação da informação é facilitada, pois, no caso do manual impresso, o usuário não precisa percorrer por várias páginas de índice para encontrar o que deseja. Portanto, bastaria realizar a pergunta como se estivesse fazendo à um especialista, como um vendedor ou mecânico.

4.3. Qual meu saldo atual? – Estudo de Caso: BIA do Bradesco

Informações sobre saldo em conta, transferência, cotação do dólar entre outras atividades financeiras fazem parte do cotidiano de muitas pessoas. Por exemplo, no caso de aposentadoria, todos que trabalham consideram a possibilidade de se aposentar um dia, sendo que, para tal feito, uma certa quantidade de dinheiro deve ser reservada para que essa aposentadoria se torne realidade (CRUZ; ALENCAR; SCHMITZ, 2018). Partindo desse pressuposto, um trabalhador é um investidor em potencial e, por conseguinte, todo investidor procura informações relevantes sobre investimentos ao seu dispor, melhores opções etc. Nesse contexto, interfaces conversacionais, como os *chatbots*, podem ajudar esses usuários a obterem não apenas informações sobre melhores investimentos, como também obterem respostas para as suas dúvidas relacionadas a sua conta bancária e serviços financeiros. Isso fica evidente com o uso do *chatbot* do Bradesco, a BIA.

A BIA (Bradesco Inteligência Artificial) é o *chatbot*, ou a assistente virtual, do Banco Bradesco (PAIVA, 2018). Por meio desse serviço, os usuários correntistas do banco conseguem obter informações financeiras por intermédio de uma simples conversa. Exemplos de informações incluem: consultar o saldo da conta; informações de abertura de conta; consultas sobre investimentos; endereço de agências. A interação com o *chatbot* é feita pelo próprio aplicativo do banco ou por meio do WhatsApp, mediante a um número disponibilizado para os clientes. Segundo o líder de Inteligência Artificial do Banco Bradesco, Marcelo Câmara (PAIVA, 2018), hoje a BIA consegue entender 94% das perguntas realizadas pelos usuários, sendo que, somente em 6% dos casos, o cliente é transferido para um ser humano.

Figura 15: Funcionamento da BIA no WhatsApp



Fonte: Print da interação com a BIA no WhatsApp

A BIA foi desenvolvida também com o uso da tecnologia do IBM Watson. Trata-se de um projeto que, inicialmente, em 2016, foi criado para atender funcionários do próprio Banco Bradesco (IBM, 2019). Com o sucesso em 2017, o sistema foi disponibilizado para o público geral. A BIA também pode ser considerada um marco na Inteligência Artificial no Brasil, pois foi por meio desse projeto que a Inteligência Artificial da IBM começou a aprender a língua portuguesa – até então, o sistema não compreendia o idioma. O *chatbot* do Bradesco é um dos poucos que se integram ao WhatsApp por meio de API. O agente é capaz de, por exemplo, utilizar o recurso de localização do aplicativo para retornar ao usuário a agência mais próxima a ele. Mediante da inteligência artificial, o agente logra compreender mais de 200 mil intenções de perguntas (IBM, 2019).

Em suma, ao utilizar a BIA, o usuário consegue obter informações de maneira bastante simples. É evidente que, muitas informações podem ser obtidas fora do *chatbot*, como por exemplo, realizar ligações diretamente para a central de atendimento do banco e falar com um atendente. Entretanto, utilizando esse tipo de agente, o indivíduo consegue obter informações rápidas, por meio de conversa em linguagem natural. Outro ponto positivo também é que não se faz necessária a instalação de qualquer tipo de aplicativo, já que a interação com o *chatbot* pode ser realizada utilizando o seu próprio WhatsApp. Devido a integração com o sistema de informação do banco, por meio de uma simples pergunta, como: “Qual o Saldo da minha conta?”, o usuário logra obter, em tempo real, a informação que deseja, evitando a necessidade

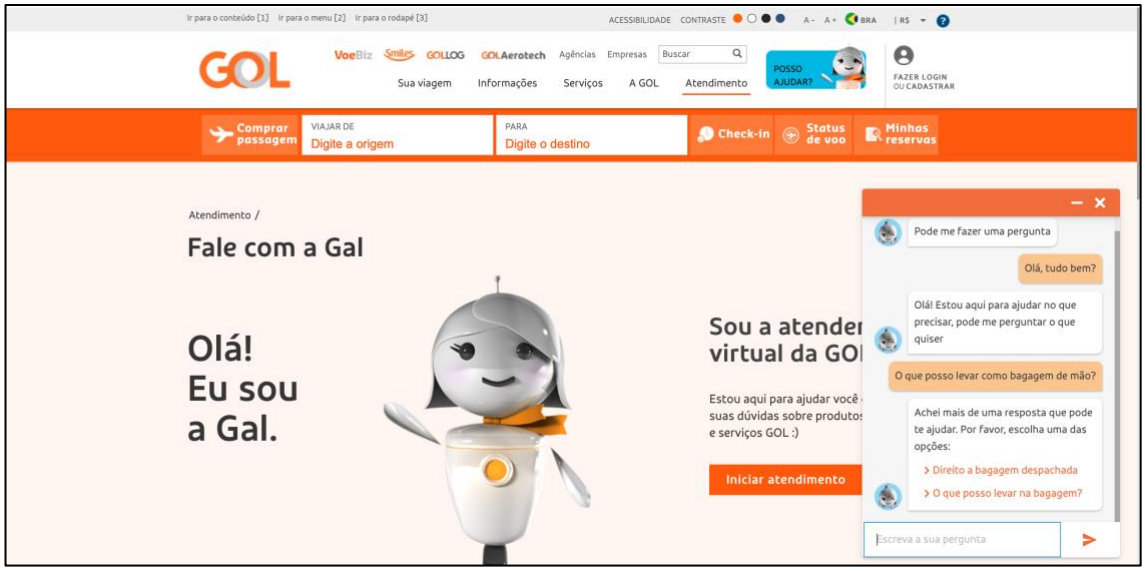
de abrir o site/aplicativo do banco, colocar usuário e senha, clicar no menu saldo para que, finalmente, obtenha a informação desejada.

4.4. Posso levar isso na bagagem? – Estudo de Caso: Gal da Gol

É conhecido que, o momento de planejar uma viagem, induz as pessoas a buscarem por diferentes informações, tais como: preços, melhores hotéis, pacotes de viagens, documentações e o que levar na bagagem. Nesse sentido, Cruz, Alencar e Schmitz (2018) sugerem que interfaces conversacionais, como o *chatbot*, podem ajudar os usuários a sanar todo tipo de dúvidas relacionadas a uma viagem, de maneira rápida e sem depender de um atendimento humano. Além disso, os *chatbots* podem auxiliar o usuário no gerenciamento de reservas, compra de seguros de viagens, indicar horários de chegada e saída dos mais variados tipos de transportes, informar sobre condições climáticas e dúvidas relacionadas ao *check-in*. Tudo isso em um atendimento focado no usuário e que está disponível 24 horas por dia. A Gal, da Gol Linhas áreas, é um exemplo nesse sentido.

Gal é o nome dado ao *chatbot* da empresa de aviação Gol Linhas Áreas. Refere-se a uma agente virtual que atende os visitantes e clientes diretamente no site da companhia. O *chatbot* fornece informações sobre bagagens; horário e informações de aeroportos; alterações de voos; entre outros. Diferente dos outros *chatbots*, toda a interação é feita por texto, e percebe-se que a Gal não tem funcionalidade de *Speech Recognition;* pelo menos não a versão disponibilizada no site. (SCHELLER, 2019) Recentemente, a Gol Linhas Áreas materializou o *chatbot* no aeroporto de Guarulhos. Em parceria com a empresa PluginBot, a companhia desenvolveu um robô físico que atende os clientes diretamente no balcão de atendimento da Gol, respondendo diversas perguntas, da mesma forma que é feito com a interação direto no site da empresa.

Figura 16: Funcionamento do *chatbot* Gal no site da Gol



Fonte: Gol Linhas Aéreas (2020)

Ao analisar o *chatbot* virtual – aquele disponibilizado no site da companhia - percebe-se que ele utiliza Inteligência Artificial, já que consegue identificar intenções por meio de perguntas digitadas em linguagem natural. Nota-se ainda, que é um *chatbot* do tipo híbrido, pois utiliza tanto Inteligência Artificial, como a interação por botões, guiando o usuário para o diálogo com as respostas mais adequadas a suas perguntas. Diferente do *chatbot* virtual, a Gal que é disponibilizada no aeroporto consegue interagir com o usuário por meio de voz e guiá-lo em diferentes pontos do aeroporto (SCHELLER, 2019), com o uso da geolocalização.

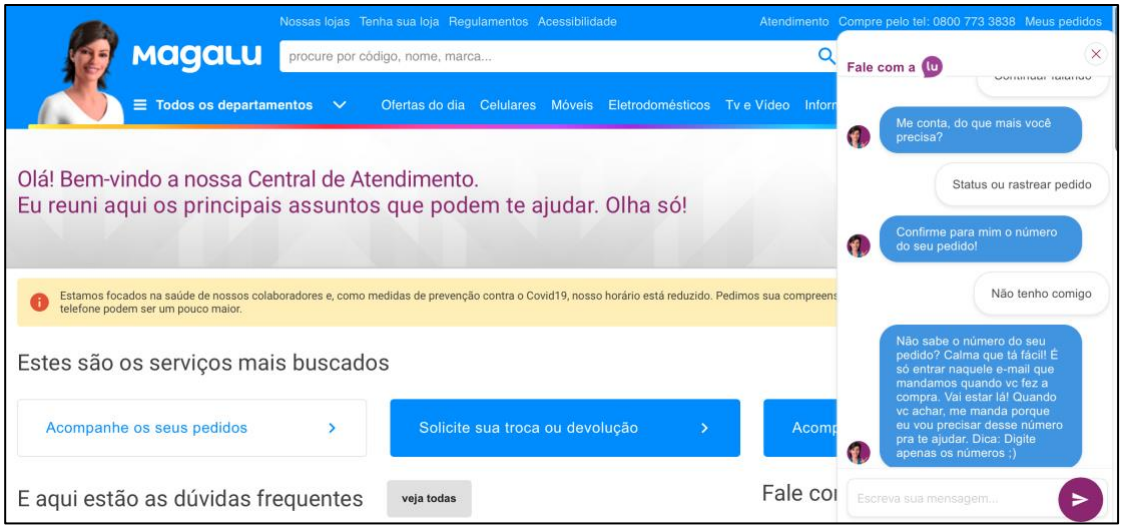
Em síntese, a GAL é um exemplo de como os *chatbots* podem ajudar o usuário no fornecimento de informações de maneira prática e interativa. Todas as informações disponibilizadas por esse *chatbot* estão disponíveis no site da companhia aérea. Portanto, o uso do agente facilita a recuperação dessa informação, já que para buscar informações sobre bagagens, o indivíduo teria que navegar pelo site, encontrar o link correspondente à aquela informação, ler todo o texto para, por fim, talvez satisfazer a sua necessidade de informação. Diferente do *chatbot*, que traz essas informações de forma resumida e são recuperadas por uma pergunta que é realizada em linguagem natural.

4.5. Aceita cartão? – Estudo de caso: Lu da Magazine Luiza

No momento de compra de produtos na internet, é comum surgir dúvidas relacionadas ao tipo de pagamento, características do produto, horário de atendimento da loja física, prazo de entrega e informações gerais sobre o produto. Essas informações são importantes para que o consumidor possa reunir um arcabouço de conhecimento para que, finalmente, efetue sua compra de forma eficaz. É sabido que o perfil do consumidor mudou (KOTLER, 2017), pois os consumidores procuram atendimento ágil e simples para que suas dúvidas sejam respondidas em tempo real, caso contrário, certamente esse indivíduo buscará outras empresas para sanar suas inquietações e, consequentemente, realizará a compra de um produto ou serviço desejado nesse outro local. É nesse contexto que os *chatbots* parecem ser uma ferramenta essencial, pois podem se configurar como um intermediador de compras, auxiliando os usuários no esclarecimento de diversas dúvidas que são sanadas em um atendimento 24 horas (CRUZ; ALENCAR; SCHMITZ, 2018). Veja o exemplo do *chatbot* Lu.

A Lu é o *chatbot* da Magazine Luiza que auxilia o usuário durante e após o processo de compra (CALADO, 2019). Trata-se de um serviço multicanal, pois a interação com ele pode ocorrer por meio do Facebook Messenger, WhatsApp, Google Assistant e diretamente no e-commerce da empresa. É um *chatbot* bastante 'simples', porém assertivo no que se propõe a fazer. Os usuários conseguem obter informações relacionadas ao status do pedido; imprimir a segunda via de boleto; solicitação de nota fiscal; busca por produtos; pedido de troca e cancelamento. Toda a comunicação é realizada em linguagem natural ou por meio de botões, guiando o usuário a resposta mais adequada à sua pergunta. A interação com o agente se dá por meio de uma comunicação bastante informal. O *chatbot* segue a mesma linguagem e personalidade da marca – descontraída e informal.

Figura 17: *Chatbot* Lu em funcionando no site da Magazine Luiza



Fonte: Atendimento Magalu (2020)

Nota-se que é um *chatbot* bastante específico e com um domínio de conhecimento restrito. Logo que uma conversa é iniciada com a Lu, automaticamente ela fornece os assuntos dos quais ela está apta a falar. Essa ação está relacionada com uma técnica para que o usuário não fuja do assunto de domínio do agente. Caso o *chatbot* não consiga compreender a informação desejada, o usuário é transferido para um atendente ou é dada a opção de telefone para que a pergunta seja esclarecida diretamente na central de atendimento da empresa. Nota-se também a integração do *chatbot* com os sistemas corporativos da empresa, como o sistema de pedidos, pois basta que o indivíduo informe o número de sua compra para que seja fornecido informações detalhadas. Tudo isso, utilizando uma interação por meio de conversas em linguagem natural.

Em suma, a integração entre sistemas de informação, como status do pedido, permite que o usuário recupere informações de maneira rápida. Tal qual, versado nos exemplos anteriores, o indivíduo poderia obter essas informações diretamente no site da empresa ou por meio da central de atendimento, seja por telefone ou e-mail. No entanto, com o uso do *chatbot*, principalmente pelo fato da interoperabilidade entre os sistemas, - o usuário consegue obter as informações sobre seu pedido de maneira instantânea, utilizando os canais de comunicação que ele já está acostumado, seja pelo Facebook Messenger, Google Assistant ou pelo próprio WhatsApp. O interessante desse *chatbot* é que ele fornece informações de forma bastante proativa. Sempre que o status do pedido é alterado, automaticamente é enviado uma mensagem

no WhatsApp do sujeito, facilitando que o indivíduo não precise recuperar essa informação de forma ativa.

4.6. Quero abrir um processo – Estudo de Caso: Judi do TJSP

Para grande parte da população, as informações relacionadas aos setores dos direitos individuais e coletivos são das mais herméticas que existem. Cruz, Alencar e Schmitz (2018) argumentam que "frequentemente, a legislação e os processos judiciais contêm termos de difícil compreensão para o cidadão comum, tais como: afecções congênitas, despacho denegatório, autarquias, aforamento, cognição exauriente, dilação probatória, controle aduaneiro, alusão etc.". Um exemplo dessa situação é o que ocorre no setor de Juizados Especiais Cíveis. Como parte das causas atendidas nesse setor não é necessário o acompanhamento de um advogado, pois muitos cidadãos, que recorrem a essas unidades, precisam retornar para as suas casas, seja porque desconhecem algum procedimento, buscam auxílios que não são de competência da Justiça do Estado ou simplesmente não levam a documentação adequada para a abertura de um processo (TJSP, 2019). É nesse contexto que um *chatbot* pode ser útil, sendo de grande importância o surgimento da Judi.

A Judi é um *chatbot* do Tribunal de Justiça do Estado de São Paulo – TJSP (2019) que auxilia o cidadão que procura esclarecer dúvidas em relação ao acesso à Justiça para resolução de causas de menor complexidade, popularmente conhecidas como pequenas causas. O agente esclarece dúvidas relacionadas a abertura de processo; documentos necessários, se precisa ou não de um advogado, endereços de Juizado Especial Cível mais próximos. Não obstante, o agente fornece modelos de petição para que o cidadão possa preenchê-los e, assim, seguir para um Juizado Especial Cível com as informações corretas para a abertura do processo desejado.

Figura 18: *Chatbot* Judi em funcionamento no site do TJSP

Fonte: TJSP (2019)

O objetivo do *chatbot* é realizar uma triagem para que o cidadão verifique se de fato sua causa deve ser tratada em algum Juizado Especial Cível e, caso positivo, o indivíduo obterá as informações necessárias para que seja dado início ao procedimento. Com o uso desse agente, o cidadão obtém esclarecimentos de forma simples e em um único lugar. O indivíduo não precisará buscar informações em diferentes fontes para satisfazer sua necessidade de informação. Todas as dúvidas são esclarecidas diretamente com o *chatbot*, em uma interação em linguagem simples e compreensiva para os indivíduos que não são da área. Desta forma, munido com essas informações, o cidadão chegará ao Juizado Especial Cível com os documentos necessários para que seja dado andamento aos devidos procedimentos.

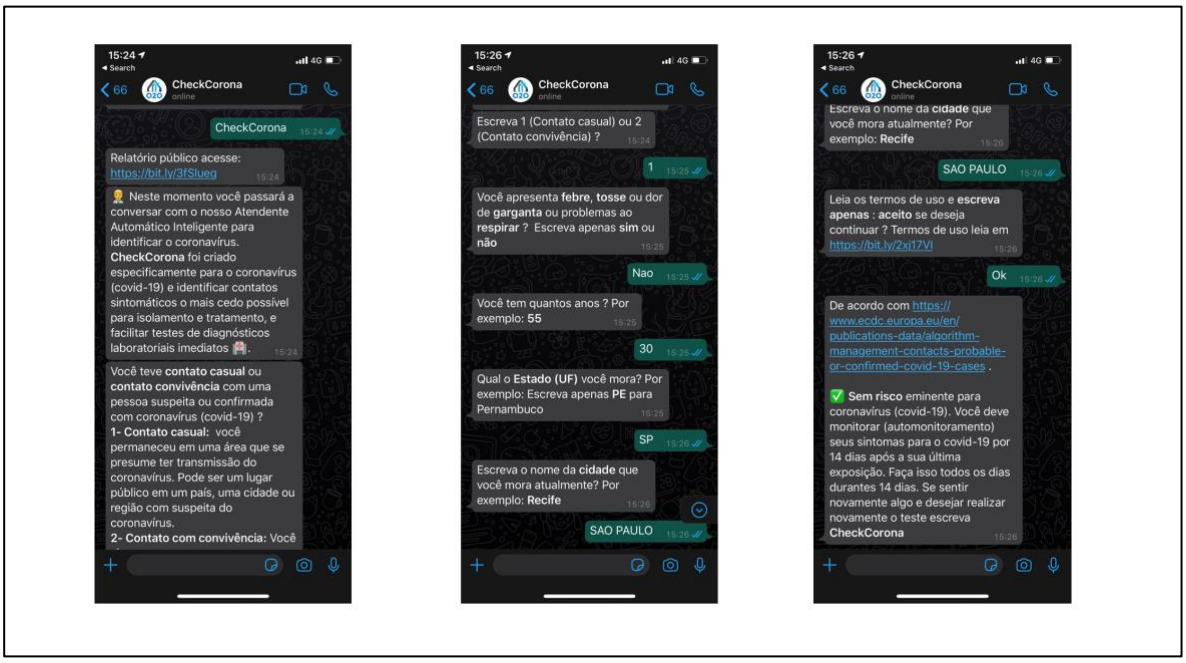
4.7. Estou com o Coronavírus? – Estudo de Caso: CheckCorona

É cada vez mais comum o uso da internet para obtenção de informações relacionadas a alguma doença. Normalmente, quando um usuário sente algum sintoma, cria-se uma necessidade de informação para saber o que aquele sintoma significa e se está relacionado a algum tipo de doença (COELHO; COELHO; CARDOSO, 2013). É sabido que a procura médica nessas ocasiões é o mais recomendado, sobretudo para evitar automedicações (COELHO; COELHO; CARDOSO, 2013). Porém, também é verdade que quando a informação é de qualidade, o usuário/paciente poderá ter um maior conhecimento de sua situação e procurar um especialista, caso seja realmente necessário. Nesse sentido, os *chatbots* podem ser um importante

intermediador nesse processo, pois são capazes de indicar diversos tipos de informações, tais como: opções de dietas, exercícios, hábitos de consumo que levam uma vida mais saudável e, obviamente, a sugestão de procurar um especialista no assunto (CRUZ; ALENCAR; SCHMITZ, 2018). Nessa área, o CheckCorona é um exemplo de serviço para tal finalidade.

O CheckCorona é o nome dado a um *chatbot* que auxilia uma pessoa a identificar se possui probabilidade de estar contaminado com o Coronavírus. O projeto foi criado em 2020 pelo Murilo Gazzola (CASATTI, 2020), doutorando do Instituto de Ciências Matemáticas e de Computação (ICMC) da USP São Carlos em plena pandemia do da COVID-19 - vírus surgido no final do ano de 2019 na China. Trata-se de um *chatbot* programado para fazer perguntas simples relacionadas aos sintomas do vírus e, com base nessas respostas, o sistema é capaz de fazer uma espécie de primeiro atendimento. Toda a interação é realizada diretamente pelo WhatsApp. Pautado nessas respostas o *chatbot* indica quais medidas o usuário deve tomar, seja ficar em isolamento ou procurar uma unidade de saúde mais próxima. Não obstante, o agente consegue agrupar dados para realizar um levantamento posterior, como quais regiões possuem maiores casos da doença (CASATTI, 2020).

Figura 19: Funcionando do CheckCorona no WhatsApp



Fonte: Print da interação com o CheckCorona no WhatsApp

Ao interagir com o CheckCorona, é possível observar que se trata de um agente que interage com o usuário por intermédio de uma conversa pré-estruturada. Todo o diálogo é realizado para que o indivíduo responda perguntas pré-determinadas para que ele prossiga em um fluxo de conversa correto e, por conseguinte, tenha um diagnóstico mais próximo da realidade. Percebe-se também o uso de Inteligência Artificial, sobretudo o Processamento de Linguagem Natural, já que o *chatbot* aceita variações de palavras em determinadas perguntas.

O projeto CheckCorona foi uma das 71 ideais selecionadas no Desafio Covid-19, iniciativa proporcionada pelo Ministério Público de Pernambuco e da Secretaria Estadual de Saúde do estado (CASATTI, 2020). Além disso, a solução foi uma das selecionadas pelo Escritório das Nações Unidas de Serviços para Projetos (UNOPS) e também pela iniciativa Dia Mundial da Criatividade para fazer parte do Mapa da Reação, que é uma plataforma online que compila projetos inovadores de respostas a pandemia. Esses reconhecimentos colaboram com a utilidade do *chatbot* no sentido de levar acesso a informação para as pessoas de maneira simples, por meio de uma conversa que se assemelha a de uma consulta médica. Mediante as informações fornecidas pelo agente, que são alimentadas a partir de normas internacionais do Centro Europeu de Prevenção e Controle das Doenças (CASATTI, 2020), as pessoas podem tomar melhores decisões e evitar a ida ao médico, caso não tenha sintomas. Essa informação também colaboraria para a não sobrecarga do sistema público e privado de saúde, além de evitar uma possível contaminação involuntária.

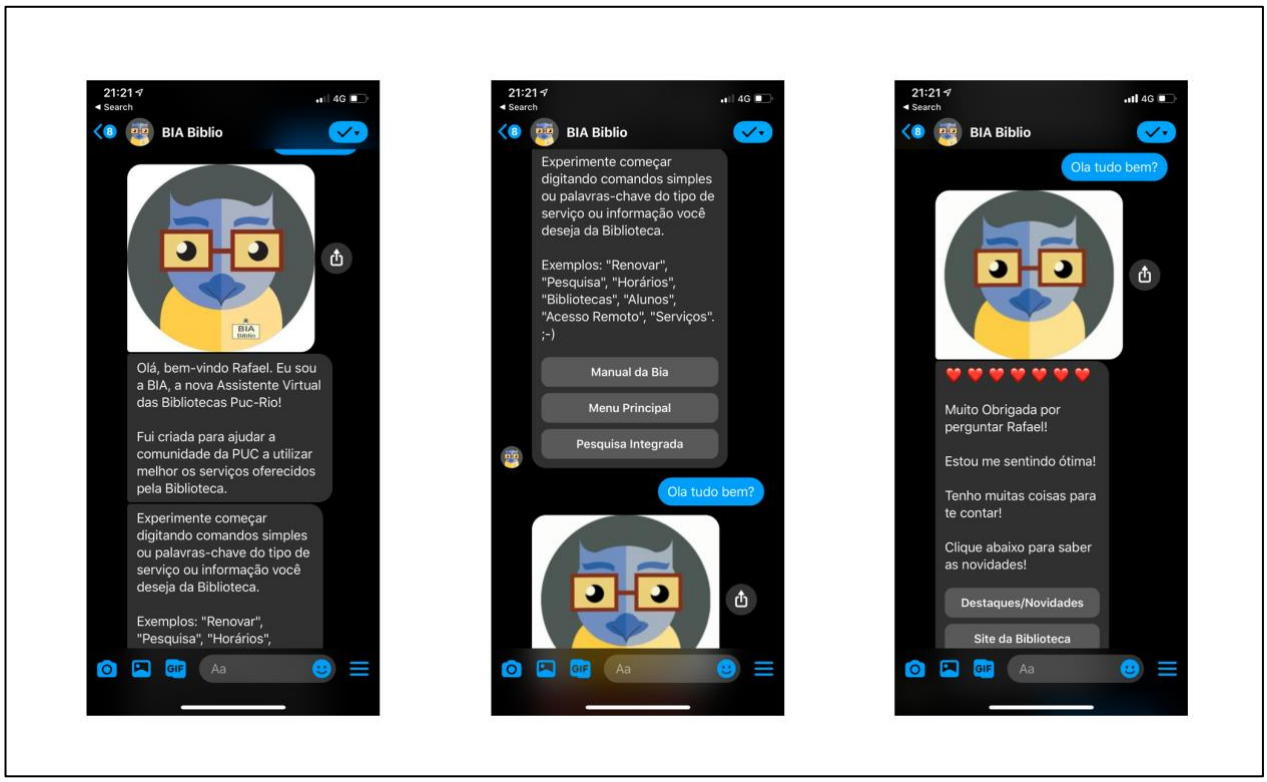
4.8. Você tem esse livro? – Estudo de Caso: Bia da Biblioteca da PUC

Em uma biblioteca é usual constar um serviço de referência para o usuário. Um serviço de referência é um processado baseado em perguntas e respostas que têm por objetivo entender e satisfazer uma necessidade de informação de um indivíduo em uma unidade de informação, como uma biblioteca (HENN, 2006). Envolve, portanto, de um lado o usuário, que lança a questão e, do outro lado, a unidade de informação que irá responder à pergunta. Nesse processo, tem-se o intermediador bibliotecário que, com sua experiência, consegue tonar a questão do usuário inteligível ao ambiente de unidade de informação (HENN, 2006). Entretanto, com a biblioteca em seu formato virtual, esses serviços de referência intermediado por um ser humano se torna cada vez mais difícil, basicamente quando se faz necessário um atendimento 24 horas. Nesse sentido, agentes inteligentes, como os *chatbots*, “podem atuar como catalisadores para

elevar o papel dos bibliotecários como a aurora do próximo século, anunciando um renascimento na ciência da informação e biblioteconomia" (HENN, 2006, apud ZICK, 2006). O pensamento de Zick parece tornar-se cada vez mais real quando se analisa *chatbots* como a Bia.

A Bia é o *chatbot* da Universidade Pontifícia Universidade Católica do Rio de Janeiro e foi criada em 2017 para simular uma bibliotecária virtual da instituição (TEBA, 2017). Na ocasião, o agente auxilia o público em diversas tarefas de serviço de referência em uma biblioteca, tais como: pesquisa e renovação de livros; horário de funcionamento da biblioteca; avisos de livros que precisam ser devolvidos. Toda a interação com o *chatbot* é feita exclusivamente utilizando o Facebook Messenger, ou seja, novamente, o usuário não precisa instalar nenhum aplicativo próprio da biblioteca para utilização do serviço.

Figura 20: Interação com a Bia via Facebook Messenger



Fonte: Print da interação com a BIA no Facebook Messenger

O *chatbot* é bastante simples, no entanto, cumpre o que se propõe a fazer: auxiliar o usuário na busca de informações relacionadas ao serviço de referência da biblioteca da universidade. Ao

interagir com o serviço, fica claro que toda sua interação é bem restrita a algumas opções de respostas. Trata-se de um agente do tipo baseado em regras, pois a comunicação se dá por meio da escolha de opções durante o diálogo. Entretanto, é possível observar que o serviço consegue entender poucas palavras escritas em texto livre, como uma simples saudação - "Olá". Todas as respostas dadas pelo agente são links que direcionam o usuário à uma página correspondente à sua pergunta. Conforme clicar na opção "Renovar", o indivíduo é transferido para a página da PUC, onde se pode realizar essa operação. Observa-se, portanto, que não existe interação nativa do *chatbot* com o sistema de biblioteca da instituição.

Em resumo, apesar de simples, é um *chatbot* que mostra sua utilidade ao auxiliar o usuário a encontrar a informação que deseja, mesmo embora toda a comunicação seja feita por um fluxo de opções por botões, e não por escrita livre em Linguagem Natural. Ou seja, é um agente que não usa recursos avançados de Inteligência Artificial. No entanto, fica evidente, com esse exemplo, que um *chatbot* pode ser utilizado como um intermediador em um serviço de referência. Não é à toa que, em 2018, o projeto recebeu, em primeiro lugar, o Prêmio de Inovação em Bibliotecas Elsevier, durante o XX Seminário Nacional de Bibliotecas Universitárias (RATTON, 2018). Isso demonstra o potencial dessa ferramenta em facilitar a recuperação de informação em um ambiente como esses; tudo isso utilizando um canal de mensagens já existente (Facebook Messenger).

Com base no que foi explanado até aqui, os *chatbots* parecem ser uma sugestão adequada de ferramenta para a recuperação da informação, tornando-se um importante mediador da informação. Isso porque, com esses sistemas, os usuários conseguem realizar buscas por informação de maneira mais natural, amigável e menos burocrática, utilizando um simples bate-papo em linguagem natural. Ao invés de navegar em um site ou manual para encontrar a informação que precisa, as pessoas simplesmente perguntam o que querem para um robô. Não obstante, as informações recuperadas tendem a ser mais eficientes, já que com a ajuda de *Machine Learning*, por exemplo, os *chatbots* são capazes de aprender com o erro e, com base nesse histórico de erros e acertos, as informações recuperadas tendem a ser exatamente as que o usuário procura.

5. RELAÇÃO E CONTRIBUIÇÃO DA CI PARA OS *CHATBOTS*

Conforme elucidado até aqui, os *chatbots* podem se figurar como uma maneira dinâmica para a recuperação da informação por meio da utilização do Processamento de Linguagem Natural. Entretanto, ao utilizar esses agentes, percebe-se algumas deficiências, por exemplo, é comum se deparar com situações onde se faz uma pergunta, porém o agente retorna com uma resposta do tipo: "Desculpe, mas não entendi o que você quis dizer". É corriqueiro o usuário mudar de assunto no meio de uma conversa e, infelizmente, muitos *chatbots* se perdem nessa troca de contexto, resultando em respostas indevidas e sem nexo ao indivíduo. Além disso, nota-se que alguns agentes não conseguem dar a mesma resposta para palavras que possuem relações semânticas próximas, mesmo embora, supostamente, aquela resposta exista na base de conhecimento do sistema.

Figura 21: *Chatbot* não reconhece a intenção de pergunta

Fonte: Submarino Viagens (2020)

A Figura 21 evidencia este tipo de situação. Percebe-se que o *chatbot* não entende o que é perguntando pelo usuário, ainda que a pergunta esteja dentro do domínio de conhecimento sobre o assunto viagens, ou seja, frases como: "reservar um voo" ou "emitir uma passagem", normalmente fazem parte do contexto de uma situação de compra de passagem aérea e possuem relações semânticas próximas ao termo "comprar". No entanto, o agente somente reconhece a intenção de compra quando o usuário envia a mensagem "Gostaria de realizar uma compra". Em outras palavras, fica evidente que o *chatbot* tem a resposta em sua base de conhecimento,

mas não é capaz de reconhecer quando o usuário faz uma pergunta que, supostamente, é diferente daquela cadastrada em sua base de conhecimento.

> Em um sistema de conversação linguístico, os seres humanos podem garantir que perguntas com o mesmo significado recebam a mesma resposta. Um sistema de aprendizado de máquina pode falhar em reconhecer corretamente perguntas semelhantes formuladas de maneiras diferentes, mesmo dentro da mesma conversa (ARTIFICIAL SOLUTIONS, 2020, tradução nossa).

Segundo a consultoria Artificial Solutions (2020), um dos principais problemas que leva o abandono de uma interação com um *chatbot* é justamente quando esse sistema não é capaz de identificar as intenções de perguntas do usuário, resultando em uma interação não satisfatória. Além disso, outro fator que corrobora para essa má experiência é a falta de informações no momento de treinar esses agentes. Sendo assim, algoritmos de *Machine Learning* e de NPL são mais eficazes quando treinados com informações organizadas (EARLEY, 2016). O que é natural para o ser humano, como entender as relações entre palavras, frases, sinônimos, entidades lexicais, conceitos etc., deve ser ensinado para uma máquina.

Nesse contexto, Seth Earley (2016) em seu artigo "*There is no AI without IA*" (Não existe Inteligência Artificial sem Arquitetura de Informação – tradução livre) argumenta que sistemas de inteligência artificial, como os *chatbots*, somente funcionam de maneira eficaz quando o trabalho de indexar, organizar, coletar e preparar a informação é feito antes, para que os algoritmos possam entendê-las e processá-las de maneira eficaz. O autor argumenta que, embora os sistemas de inteligência artificial estejam recebendo bastante visibilidade por sua promessa de facilitar a recuperação da informação, o fato de requerer um trabalho prévio de organização do conhecimento e representação da informação é pouco conhecido pelos profissionais que desenvolvem esses sistemas e, muitas das vezes, essa etapa é ignorada.

> Muitos fornecedores contornam essa questão ou afirmam que seus algoritmos operam em fontes de informação não estruturadas, "entendem" essas fontes, interpretam a consulta do usuário e apresentam o resultado sem arquiteturas predefinidas ou intervenção humana. Isso pode muito bem ser verdade em determinadas circunstâncias, mas a maioria das aplicações exige uma quantidade significativa de trabalho árduo por parte dos humanos antes que as redes neurais, o aprendizado de máquina e os processadores de linguagem natural possam trabalhar sua mágica – tradução do autor (EARLEY, 2016)

Portanto, apesar dos sistemas de inteligência artificial serem uma promessa no que tange a automação de processos e recuperação da informação, fica evidente que eles não funcionarão de maneira apropriada caso não exista um trabalho preliminar para ensinar essas máquinas. "Se a informação não está estruturada e curada de alguma maneira, não há nada que os sistemas

possam encontrar" (EARLEY, 2016). A Ciência da Informação, por meio da área da Organização da Informação e do Conhecimento, pode colaborar muito nesse processo.

O termo Organização do Conhecimento foi cunhado pela primeira vez em 1971 por Henry Evelyn Bliss (ALVARES, 2012) e trata-se da área de estudos voltada para atividade de organização, representação e recuperação da informação. Dahlberg (ALVARES, 2012) define essa área como uma ciência que estrutura e organiza de maneira sistemática unidades de conhecimento (conceitos) segundo seus elementos de conhecimento (características) que são inerentes a aplicação de conceitos e classes de conceitos ordenadas por objetivos ou assuntos. É, portanto, uma etapa fundamental para que sistemas de recuperação da informação, como *chatbots*, funcionem de maneira adequada. Isso porque, ao passo que a informação está bem estruturada, os sistemas de informação são capazes de inferir de forma mais eficaz com base nas entradas de busca ou perguntas do usuário.

> Em um sistema de informação, a qualidade obtida na recuperação da informação depende substancialmente dos procedimentos e instrumentos utilizados na organização da informação. Os padrões de organização devem, portanto, ser definidos desde a concepção do sistema, para permitir que a informação seja encontrada posteriormente. (ALVARES, 2012)

Em suma, percebe-se que os problemas de interação dos *chatbots* estão relacionados ao desenvolvimento de um fluxo de diálogo coerente; a uma melhor organização de entidades e intenções; um melhor entendimento do perfil dos usuários; a compreensão correta do domínio e vocabulário utilizado pelo *chatbot*; e, finalmente, algum método para entender as reais necessidades de informação e comportamento de busca do usuário. Nesse contexto, a Ciência da Informação, mais especificamente a Organização do Conhecimento, possui ferramentas que podem contribuir com a resolução desses problemas e com o processo de desenvolvimento e melhoria de um *chatbot*, tendo como resultado final, um agente que compreenda e entenda melhor o usuário no momento de sua interação com a finalidade de recuperação da informação.

A seguir, relata-se como a Ciência da Informação pode colaborar para o desenvolvimento e o processo de melhoraria de um *chatbot*, ressaltando que não se pretende esgotar os temas que serão abordados aqui. O intuito é mostrar algumas relações e ferramentas que vem sendo estudas dentro da CI, há bastante tempo, e que podem ser úteis para melhorar e mitigar erros de interação com um *chatbot*. Com isso, em cada um dos erros a seguir, serão utilizados diálogos fictícios para exemplificar algumas falhas de interação que podem ocorrer quando não há um preparo da informação nesses sistemas. Por não ter sido encontrado na literatura essas falhas

evidenciadas de maneira clara, os diálogos fictícios foram criados com base na utilização de alguns *chatbots* e conhecimento empírico deste autor.

5.1. Estudos de usuário da informação para definição de persona

Conforme visto no decorrer desta pesquisa, persona é o nome dado a um conjunto de características de um *chatbot*, isto é, a linguagem utilizada, perguntas que está apto a responder, contexto das conversas, tom e mais. Conhecer o usuário é uma das etapas essenciais no desenvolvimento de um sistema de recuperação da informação, tal como um *chatbot* (BAPTISTA; CUNHA, 2007). Portanto, é a partir das características e anseios desses indivíduos que todos os passos de construção do agente serão pautados. Isso inclui: tipo de linguagem, vocabulário utilizado, tipo de perguntas, a definição de persona para o agente, entre outros. O conhecimento sobre o usuário é importante, pois este é o responsável pela a existência de um serviço e, por conseguinte, por avaliá-lo a partir do seu uso (RIBEIRO, 2020). Observe o diálogo hipotético com um *chatbot* que fornece informações jurídicas a um usuário comum:

Usuário: Oi, gostaria de saber como abro um processo.
Chatbot: Excelentíssimo, por obsecro, elucide sobre qual tipo de processo deseja informações.
Usuário: Obsecro? Enfim, quero processar um e-commerce.
Chatbot: Prezado, você quer informações de como iniciar uma exordial?
Usuário: O que é exordial? Não estou entendendo nada.

O diálogo evidencia o que acontece quando um projeto de *chatbot* não tem claro o perfil dos usuários que o utilizarão; tão pouco seus anseios, ideias, cultura e contexto. Nesse caso, por não compreender o usuário, o agente utiliza uma forma de comunicação - conhecida como juridiquês – que não pertence ao contexto de uma pessoa comum. Não obstante, percebe-se um tom extremamente formal, o que pode afastar ou atrapalhar o usuário no momento da sua interação. Nesse sentido, a Ciência da Informação, por meio dos estudos de usuário, tem muito o que colaborar para que os profissionais envolvidos no desenvolvimento de um *chatbot* consigam, de fato, conhecer e compreender as características e necessidades de informação do indivíduo que irá utilizar o serviço. No âmbito da Ciência da Informação, o conceito usuário pode ser definido como:

> O conceito de usuário é complexo. Diz respeito tanto ao especialista que interroga uma base de dados como aquele que solicita um serviço (resposta a uma questão ou uma pesquisa bibliográfica); ao cliente de um serviço; ao produtor de informação; entre outros. (DIAS, 2004)

Ferneda (2003) lembra que "[…] os sistemas de informação deveriam ser modelados conforme a natureza das necessidades de informação do usuário, levando-se em conta os seus padrões de comportamento na busca da informação". Portanto, o usuário é um elemento fundamental e conhecer seu contexto e suas necessidades de informação, passam a ser tarefas essenciais para o desenvolvimento de um Sistema de Recuperação da Informação (SRI) eficiente, como um *chatbot*.

Entende-se por necessidade de informação o anseio que faz com o que o usuário busque por determinado tipo de informação, para auxiliar na resolução de algum problema (FERNEDA, 2003; SOUZA, 2006). Na CI, a necessidade de informação está representada pela expressão de busca (termo que o usuário insere para buscar determinada informação). Essa expressão de busca pode ser em linguagem natural (aquela utilizada convencionalmente pelas pessoas) ou por meio de uma linguagem artificial, do tipo tesauro ou vocabulário controlado, elaborada especificamente para representação e recuperação da informação em um SRI". (FERNEDA, 2003; SOUZA, 2006). Dias (2004) argumenta que compreender as necessidades de cada indivíduo em relação a informação é algo bastante complexo e que se modifica constantemente, o que reforça a necessidade de contextualizar os estudos de usuários realizados na CI como uma contribuição ao desenvolvimento dos *chatbots*.

Os estudos relativos aos usuários da informação datam o ano 1930, onde naquela ocasião, a Universidade de Chicago realizou os primeiros estudos com usuários de bibliotecas com o objetivo de identificar hábitos de leitura e como eles utilizam os acervos disponíveis (ARAÚJO, 2012). Entretanto, em 1948, na Conferência da Royal Society, foram apresentados trabalhos que contribuíram para o crescimento de estudo voltados para as necessidades do usuário (FIGUEIREDO, 1994). Em 1970, os estudos de usuários da informação direcionaram-se para o viés de entender o indivíduo como ser humano, analisando seus anseios, comportamento de busca, cultura, e, sobretudo, focados em pesquisas orientadas a satisfação de suas necessidades de informação (DIAS, 2004). Figueiredo (1994, p.7) define estudos de usuário da seguinte maneira,

> Estudo de usuários são investigações que se fazem para saber o que os indivíduos

> precisam em matéria de informação, ou então para saber se as necessidades de informação por parte dos usuários de uma biblioteca ou um centro de informação estão sendo satisfeitas de maneira adequada.

Segundo Cunha (1982) e Figueiredo (1994), no decorrer de sua história, os estudos de usuários da informação foram classificados em dois tipos (ou abordagens): quantitativa (abordagem tradicional) e qualitativa (abordagem alternativa). A primeira refere-se à investigação de como as bibliotecas ou centro de informações são utilizados pelo usuário. Isto é, esses estudos se preocupavam em identificar, por exemplo, a frequência que um acervo de uma biblioteca era utilizado, entre outros comportamentos que traziam resultados puramente quantitativos. Já o segundo tipo de estudo, preocupava-se fundamentalmente com o usuário como indivíduo, ou seja, investiga como o usuário obtém informação, seus anseios, necessidades de buscas e contexto.

> Passou-se de uma fase quantitativa para uma fase qualitativa quando os estudiosos do comportamento de busca da informação perceberam que as pesquisas com métodos quantitativos não contribuíam para a identificação das necessidades individuais e para a implementação de sistemas de informação adequados a essas necessidades. (BAPTISTA; CUNHA, 2007)

Nesse sentido, a abordagem alternativa parece ser mais bem aplicada como metodologia para entender o usuário, já que no processo de desenvolvimento de um *chatbot*, o objetivo é analisar como esses estudos podem ajudar no desenvolvimento e melhorias de um agente sob o prisma de aprimorar a sua recuperação da informação. A seguir, de forma sucinta, alguns exemplos de metodologias dos estudos de usuário da informação que podem ser incorporadas no processo de desenvolvimento de um *chatbot*. Dentro da abordagem alternativa, existem diversos métodos focados nas necessidades de informação do usuário, sendo alguns deles: estudos de Taylor (1982, p. 342 apud BAPTISTA; CUNHA, 2007), o modelo de Kuhlthau (1999, apud BAPTISTA; CUNHA, 2007) e, principalmente, o *Sense Making* de Brenda Dervin (1998 apud BAPTISTA; CUNHA, 2007) – que será o foco neste estudo.

A abordagem ou o modelo de informação com valor agregado de Taylor (1982, p. 342 apud BAPTISTA; CUNHA, 2007) analisa a questão da busca da informação pelo usuário e o processo de transformação de dados em informação útil.

> Para o autor, os processos de seleção, análise e julgamento podem transformar um dado em informação útil. Essa informação poderá ser empregada para esclarecer, informar e contribuir em relação ao crescimento pessoal, cultural e afetar as decisões e ações pessoais do usuário de um sistema de informação. (BAPTISTA; CUNHA, 2007)

Essa abordagem reúne métodos que fornecem insumos de como transformar um dado (ou informação sem muito sentido) em algo útil para o usuário. Por meio de processos como organização, análise, síntese e julgamento (KIELGAST; HUBBARD, 1997), essa abordagem pode auxiliar na análise de um *chatbot*, no sentido de verificar a qualidade da informação fornecida, tal como foi feito nos estudos de Souza e Borges (1999 apud BAPTISTA; CUNHA, 2007) ao analisar a qualidade de fontes de informações financeiras produzidas no Brasil.

Já o modelo de Kuhlthau (1999, apud BAPTISTA; CUNHA, 2007) analisa o comportamento do usuário na busca de uma informação. Esse modelo foi desenvolvido após pesquisas sobre a experiência de usuários em bibliotecas e em ambientes de trabalho.

> Para a autora, as pessoas que buscam informação utilizam muitas fontes de informação em diversas fases da busca e, esse processo gera incerteza para o usuário em distintos momentos da sua busca. Seu modelo para observação do processo da busca da informação prevê as seguintes tapas: início, seleção, exploração e formulação. (KUHLTHAU, 1999 apud BAPTISTA; CUNHA, 2007)

Trata-se de um modelo que analisa todas as fases em que o usuário busca pela informação. Sendo assim, a fase inicial (início) ocorre quando o usuário sente falta de alguma informação para solucionar um problema. A segunda fase (seleção) é o momento no qual o usuário seleciona algumas informações que podem ser úteis para a resolução do seu problema. Já a fase seguinte (formulação) é o momento onde, diante das informações selecionadas, o usuário consegue explorar e satisfazer sua necessidade de informação. Em síntese, ao colocar o usuário como agente principal desse processo, esse modelo permite aperfeiçoar o desenvolvimento de sistemas de recuperação da informação mais eficientes, como os *chatbots*.

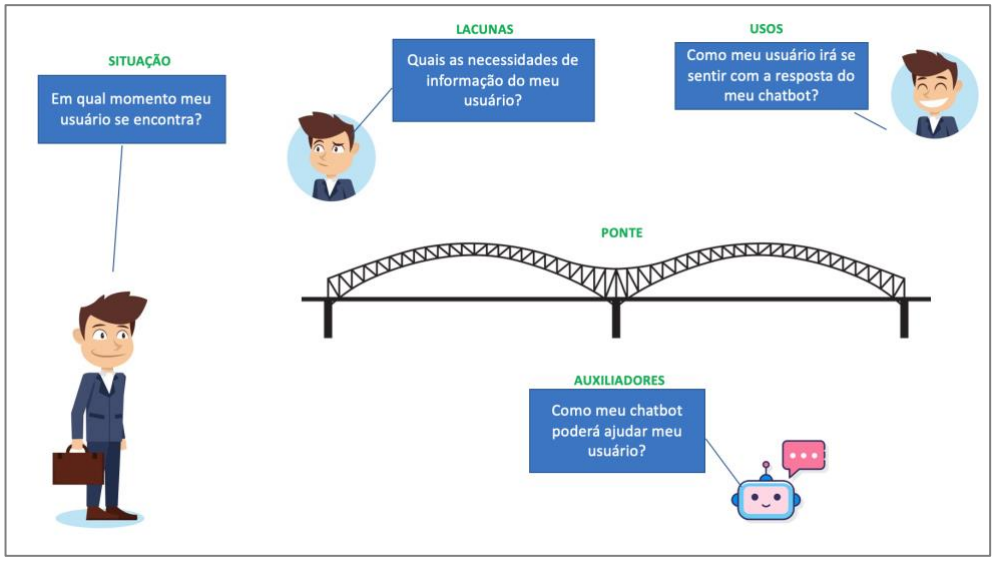
Por fim, e como método de destaque para a presente pesquisa, a metodologia *Sense Making* de Brenda Dervin (1998 apud BAPTISTA; CUNHA, 2007) diz respeito a uma abordagem a qual promove uma forma de entender as necessidades de informação dos usuários, sob um viés cognitivo, afetivo, psicológica e fisiológica.

> A metodologia *Sense-Making* foi desenvolvida pela autora desde 1972, ganhando uma formulação definitiva por volta do início da década de 1980. Essa abordagem, que como o nome diz, literalmente significa “fazendo sentido”, e consiste em um método, uma forma de compreender a necessidade de informação do usuário e sua relação com o acesso e uso da mesma. (NOGUEIRA; GRACIOSO; TRIQUES; VIEIRA, 2015)

É um método que analisa, de forma bem detalhada, a necessidade de informação do usuário. Bax & Dias (1997) dizem que existem quatro pontos fundamentais vividos pelo ser humano em

uma situação de busca. Esses pontos, segundo os autores, podem ser explicados na metáfora de uma pessoa atravessando uma ponte. Os pontos são os seguintes: situação, lacuna, auxiliadores e a ponte. A situação é o ponto que caracteriza por um conjunto de ideias, influências, restrições, contexto e experiências do usuário. Já a lacuna, pode ser entendida como as questões, angústias e as necessidades de informação do usuário. Os auxiliadores são ideias e conceitos que ajudarão o usuário a satisfazer as suas necessidades de informação. Por fim, a ponte, que são as ideias formadas por meio das informações obtidas através dos auxiliadores e que levarão o usuário ao uso do conhecimento (DERVIN, 1983) e, consequentemente, a solução do problema. A Ilustração 12 é uma adaptação das ideias de Bax & Dias (1997) e Dervin (1983).

Ilustração 12: Aplicação da metodologia de *Sense making* para os *chatbots*



Fonte: O autor (2020) – Adaptação de Bax & Dias (1997) e Dervin (1983).

A Ilustração 12 contém algumas sugestões de indagações que o profissional responsável por desenvolver um *chatbot* poderia fazer a si mesmo ou diretamente ao usuário, tendo como intuito conhecê-lo melhor ou, utilizando a analogia da ponte, ajudá-lo a atravessar a mesma. A "**situação**", seria o momento de entender qual seria o contexto que levaria o usuário a utilizar o *chatbot*, além do seu perfil, cultura, conhecimentos prévios e linguagem. Já as "**lacunas**", são os principais problemas e necessidades de informação que esse usuário teria no momento de interagir com o agente. Os "**auxiliadores**" representam as reflexões sobre como o sistema pode auxiliar esse usuário, considerando sua situação e contexto. Por fim, os "**usos**" são as indagações de como o usuário se sentiria ao receber a informação e como os seus problemas seriam resolvidos com ela.

É evidente que compreender o usuário, seu comportamento, anseios, necessidades e contexto, são passos fundamentais para o desenvolvimento de serviços e sistemas de recuperação da informação mais eficientes, como um *chatbot*. A distribuição massiva de informação não garante a qualidade e melhoria de vida das pessoas. É necessário pensar como essa informação irá contribuir para que o sujeito processe e a transforme em conhecimento (SMIT, 2012). A partir do momento que o profissional envolvido no processo de construção de um *chatbot* desenvolve o seu trabalho questionando como o usuário recuperará e processará a informação, tem-se a probabilidade de construir um SRI eficiente e que, por conseguinte, deixará o usuário satisfeito.

> Nos dias atuais é relevante considerar que sem informação acessível e organizada é impossível exercer cidadania. Portanto, ser um curador da informação se tornou requisito básico na sociedade em que vivemos. Devemos pensar informação e o acesso a ela pelo viés de cidadão e consumidor. (RIBEIRO, 2020)

Em suma, os profissionais envolvidos no desenvolvimento de um *chatbot* poderão utilizar essas técnicas de estudos para entender melhor o comportamento de seus usuários e, como resultado, criar uma personalidade mais fidedigna ao público que irá interagir com o agente. Além disso, com informações mais detalhadas sobre as pessoas que utilizarão o agente, tem-se a possibilidade de criar uma persona que dialogue de maneira mais eficaz, baseando-se nas características, anseios e o tipo de linguagem dos usuários. Sendo assim, uma melhor compreensão do usuário pode facilitar o processo de identificar as intenções e entidades apropriadas para que se tenha um fluxo de diálogo coerente e que atenda as expectativas do indivíduo.

5.2. A indexação para identificar entidades e intenções

Depois de tudo o que foi abordado até aqui, fica evidente que uma das etapas fundamentais no processo de desenvolvimento de um *chatbot* é o levantamento de intenções e entidades, isto é, quais serão as perguntas que o agente será capaz de responder ao usuário. Como visto anteriormente, intenção é o que o usuário deseja que o *chatbot* faça, seja esclarecer uma dúvida ou realizar alguma ação. Já as entidades, são os elementos que irão complementar essa intenção. No entanto, saber identificar essas intenções e entidades torna-se um trabalho essencial para garantir uma boa recuperação de informação. A falta de exaustividade na coleta dessas

informações, faz com que o *chatbot* tenha pouca informação a fornecer, mesmo ele sendo um especialista em um determinado assunto. Veja o exemplo hipotético:

Usuário: Olá, tudo bem?
Chatbot: Olá, seja bem-vindo. Sou especialista na história do Brasil. Como posso te ajudar?
Usuário: Quem descobriu o Brasil?
Chatbot: Eu ainda não tenho a resposta para essa pergunta.
Usuário: :-0

Talvez o exemplo seja algo caricato, entretanto, porém evidencia que erros como este podem acontecer em outros contextos. Embora o *chatbot* seja especialista em uma determinada área do conhecimento, a não identificação de perguntas simples evidencia a má qualidade no processo de levantamento de intenções e entidades para a construção de sua base de conhecimento. Às vezes, o próprio site da instituição, em sua sessão de perguntas e respostas, já indica a solução para a pergunta em questão. Contudo, por uma curadoria ineficaz, perguntas simplesmente podem não ser consideradas na base de conhecimento de um *chatbot*, causando uma interação insatisfatória para o usuário e, consequentemente, não atendendo as suas necessidades de informação.

O processo de identificar intenções e entidades é feito por meio da curadoria de conteúdo (BOTS BRASIL, 2020). Os profissionais envolvidos no desenvolvimento de um *chatbot* utilizam diversas fontes de informações, tais como: website da empresa; histórico de atendimento; análise de e-mails; conversas telefônicas, entre outros. Essas fontes são utilizadas com o objetivo de encontrar quais são as principais perguntas e necessidades de informação do usuário sobre determinado tema ou empresa. Por exemplo, ao desenvolver um *chatbot* para atender clientes de uma agência de turismo, convém analisar, seja pelo site da empresa ou histórico de atendimento, quais foram as principais perguntas ou dúvidas (intenções/entidades) que o usuário fez sobre os serviços oferecidos por essa agência. Exemplo: “Qual o preço da passagem para São Paulo?” ou “Qual o horário de atendimento da agência”. Com base nessas duas perguntas, é possível identificar 2 intenções: uma relacionada a cotação de preços de passagens e a outra relacionada ao horário de atendimento da agência.

Esse processo de identificar perguntas (intenções e entidades) em um documento é bastante similar com o processo de indexação estudado a bastante tempo dentro da Ciência da

Informação, onde o intuito é identificar **conceitos** e nomeá-los por meio de **termos**. Dessa forma, trazendo para o contexto do *chatbot*, seria o mesmo processo de identificar **perguntas** e agrupá-las em **intenções**. Indexação é, portanto, um processo importante para a eficiência de um sistema de recuperação da informação (ALVARES, 2012), como um *chatbot*. Cardoso Filho e Santos (2012) definem a indexação como:

> Uma maneira de organizar a informação é a indexação, técnica de análise documentária na qual a informação significativa de um documento é traduzida com a atribuição de termos selecionados da linguagem natural (utilizada nas necessidades da vida diária) ou por símbolos, os quais intermediarão a comunicação entre o usuário e o documento. Trata-se, portanto, de um tipo de representação do conhecimento registrado.

A indexação é um processo de organizar a informação com o intuito de recuperá-la. Isso significa traduzir o que está em um documento para uma linguagem mais acessível e padronizada. É um processo que tem por objetivo resumir um documento por meio da identificação de termos que o represente. Essa sumarização facilita a busca e a recuperação da informação pelo usuário.

Segundo Fujita (2003), a indexação é uma prática bastante antiga no tratamento de documentos. Em bibliotecas da antiguidade já existiam listas de documentos que pertenciam a essas instituições. Porém, a partir do momento que essas listas precisaram ser organizadas por assuntos, e não mais pelo nome das obras, necessitou-se de um processo humano de análise de conteúdo mais apurada. O processo de indexação de documentos tornou-se ainda mais evidente com o aumento da publicação de periódicos e da literatura técnica científica impulsionando a necessidade da criação de mecanismos de controle bibliográfico em centros de documentação especializados (FUJITA, 2003).

Indexador é o nome dado ao profissional responsável por intermediar e analisar um documento a ser indexado, convertendo a informação primária em secundária, como catálogos e índices de bases de dados. Além disso, conforme elucidam Cardoso e Santos (2012), compreender o usuário é uma etapa fundamental nesse processo:

> A eficiente indexação de assuntos precisa levar em conta o usuário final. O mesmo documento pode ser indexado de maneiras diferentes, de acordo com o público ao qual se destina. Assim, torna-se recomendável que o indexador desempenhe outras atividades, como por exemplo, bibliotecário de referência, quando estará em contato direto com as necessidades da comunidade a ser atendida. (CARDOSO; SANTOS, 2012)

O bom ou mal desempenho de indexação reflete-se na recuperação da informação pelo usuário. Por esse motivo, o indexador deve ser um professional bastante capacitado, pois é ele o responsável por examinar o texto, identificar conceitos, abordá-lo de forma lógica a fim de selecionar termos que melhor representem o conteúdo de um documento (FUJITA, 2003). Logo, perguntas como "Do que se trata o documento?", "Por que foi incorporado ao meu acervo?" "Quais são os aspectos de interesse dos usuários?" são algumas questões que nortearam o indexador no processo de analisar e identificar termos que representem melhor o documento a ser indexado.

Segundo Cardoso Filho e Santos (2012) a indexação pode ser manual ou automática. A indexação automática pode ser descrita como um processo de identificação de palavras relevantes (descritores) com a ajuda de recursos tecnológicos, como ferramentas oriundas da inteligência artificial. Já a indexação manual, é uma indexação baseada em conceitos e realizada pelo homem. Na indexação manual, tem-se a possibilidade da utilização de técnicas, como vocabulários controlados, que auxiliam na precisão da recuperação da informação, tais como: controle de sinônimos, diferenciação de homógrafos e reunião ou ligação de termos associados. A presente pesquisa ressalta a indexação manual por apresentar maior riqueza de detalhes e procedimentos que podem auxiliar no desenvolvimento e melhorias de um *chatbot*.

Nesse sentido, Filho Cardoso e Santos (2012; apud FUJITA, 2003) também versam que o processo de indexação manual possui três fases fundamentais:

> [...] (i) a compreensão do conteúdo do documento, com o exame e estabelecimento do seu conteúdo, o que se realiza mediante a leitura do título, da introdução e das primeiras frases de capítulos e parágrafos, das ilustrações e tabelas, dos diagramas e suas explicações, da conclusão, de palavras ou grupos de palavras sublinhadas ou impressas com tipo diferente; (ii) identificação de conceitos, tarefa desenvolvida após o exame do texto para, de maneira lógica, selecionar os conceitos que melhor representem o seu conteúdo; e (iii) a tradução desses conceitos em linguagem de indexação, dotando-os de significado para o usuário a fim de que possam ser recuperados segundo suas necessidades

A primeira fase corresponde a compreensão do conteúdo pelo indexador. O método utilizado nessa etapa é a Análise de Conteúdo (ou também conhecido como Análise de Assunto). Trata-se de técnicas que auxiliam o indexador na descrição do conteúdo das mensagens, permitindo traduzir a informação e o conhecimento a elas associadas (FILHO CARDOSO; SANTOS, 2012). Essa análise é aplicada sob a linguagem verbal, bem como a imagens, desenhos, pinturas, cartazes, entre outros tipos de comunicação não verbal. Em síntese, a análise de conteúdo

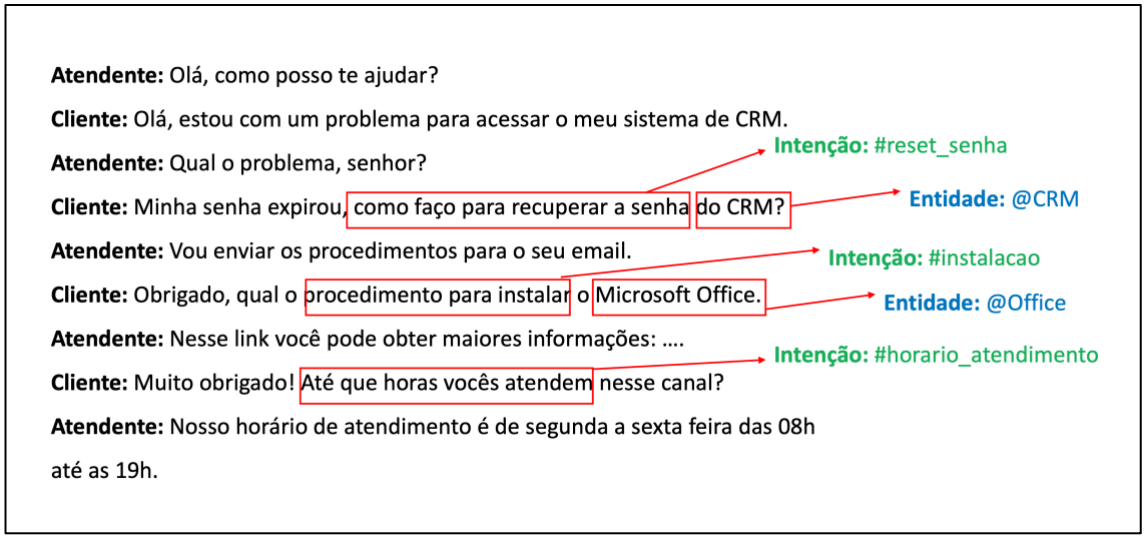
permite que o indexador busque a informação que existe além da informação manifestada (FILHO CARDOSO; SANTOS, 2012), facilitando a identificação de conceitos que atendam a necessidade de busca do usuário.

Já a segunda fase, após a compreensão do conteúdo, busca-se pela identificação de conceitos (ou termos). Aqui, após o exame do conteúdo, o indexador passa a abordar o texto de forma lógica, a fim de identificar termos que melhor representem o conteúdo. Segundo Fujita (2003), é recomendado que "a identificação de conceitos seja feita obedecendo a um esquema de categorias existente na área coberta pelo documento, como por ex.: o fenômeno, o processo, as propriedades, as operações, o material, o equipamento etc." Nessa fase, o objetivo é a extração do maior número possível de termos que possam cobrir o conteúdo de maneira mais completa. Após uma segunda análise, ficam os termos que mais representam o documento em questão.

Portanto, a terceira fase trata-se da seleção e tradução dos termos selecionados de acordo com uma linguagem de indexação, tais como vocabulário controlado, listas de cabeçalhos de assuntos e esquemas de classificação. Nem todos os conceitos identificados na segunda fase serão necessariamente utilizados. Isso é importante, pois ao contrário da exaustividade utilizada na segunda fase, a especificidade se torna fundamental para aumentar a precisão na recuperação da informação, uma vez que considera a intensidade de um termo em relação ao conteúdo de um documento (CARDOSO FILHO; SANTOS, 2012). O uso de uma linguagem de indexação torna a lista de termos mais precisa e aderente ao conteúdo a ser indexado.

O procedimento de indexação oriundo da Ciência da Informação pode ser bastante útil para o processo de curadoria, ou levantamento de entidades e intenções em um projeto de desenvolvimento de um *chatbot*. A Ilustração 13 elucida como o método de indexação poderia ser aplicado em um documento de histórico de atendimento.

Ilustração 13: Levantamento de entidades e intenções no log de um atendimento



Fonte: O Autor (2020)

A Ilustração 13 é apenas um exemplo fictício de um histórico de atendimento relacionado ao suporte de TI de uma empresa. Veja que por meio de uma simples conversa, é possível analisar o documento e extrair possíveis entidades e intenções para serem utilizadas em um fluxo de conversa dentro de um *chatbot*. Observe que o processo é bastante semelhante ao adotado no processo de indexação de documentos realizado na Ciência da Informação. Portanto, todas as etapas e procedimentos utilizados pela indexação, podem ser adotadas no momento de identificar perguntas e conteúdos relevantes para alimentar a base de conhecimento de um agente.

Fica evidente que a indexação é um processo muito importante para o desenvolvimento de um sistema de recuperação da informação, já que esse procedimento constrói a base essencial para que os usuários e os sistemas possam encontrar a informação desejada. Ao analisar as três etapas, nota-se também que a análise de assunto é uma das etapas mais importantes desse processo, já que identifica conceitos que representam a essência de um documento. Portanto, refere-se a um processo que exige bastante esforço do indexador, pois o obriga a seguir uma metodologia adequada a fim de obter resultados satisfatórios no processo de recuperação.

Em resumo, a indexação é um procedimento realizado no tratamento informacional, mais especificamente para garantir a precisão e pertinência na representação e recuperação da informação de acordo com as necessidades dos usuários e, por esse motivo, pode contribuir no processo de identificação de intenções e entidades que, posteriormente, serão utilizadas em uma

plataforma para o desenvolvimento de um *chatbot*. Analisando os documentos, como o histórico de atendimento e as mensagens trocadas entre o usuário e o *chatbot*, os profissionais responsáveis por desenvolver o agente, poderão identificar quais são as perguntas mais frequentes e agrupar essas perguntas em intenções; tal como é feito no processo de indexação, onde a primeira etapa é identificar os conceitos, para depois nomeá-los por meio de termos que organizados e definidos, constituem uma linguagem de indexação, a ser utilizada tanto no momento da busca quanto no momento da representação da informação.

5.3. O uso dos tesauros na padronização de intenções e entidades

Conforme versado no começo deste capítulo, um dos principais problemas apresentados por um *chatbot* é o fato dele não relacionar o uso de diferentes termos para uma mesma resposta. Após o levantamento de termos (ou entidades e intenções, no contexto dos *chatbots*), se faz necessário a padronização desse vocabulário. Isso tende a melhorar a interação com o agente, visto que ele poderá ser capaz de fornecer informações sobre perguntas que podem ser feitas de diferentes maneiras, mitigando respostas como: "Desculpe, mas não entendi a sua pergunta". Considere como exemplo o diálogo fictício:

Usuário: Como faço para comprar um refrigerador?
Chatbot: Entendi que você quer comprar algo, mas não entendi o que.
Usuário: Gostaria de comprar um novo refrigerador, pois o meu atualmente está velho.
Chatbot: Continuo não entendendo. Pode ser mais específico?
Usuário: Quero comprar uma geladeira!!!!!!!
Chatbot: Ok, entendi. Qual marca você deseja?

Veja que o *chatbot* não reconhece que a pergunta "*Como faço para comprar um* ***refrigerador***?" possui o mesmo significado de "*Como faço para comprar uma* ***geladeira****?"* Observa-se que a intenção é a mesma "comprar algo" e, isso foi identificado pelo agente em sua resposta "*Entendi que você quer comprar algo, mas não entendi o que.* ". Embora o *chatbot* tenha uma resposta para a pergunta em sua base de conhecimento, o fato de o usuário trocar o termo (ou entidade) **geladeira** para **refrigerador** fez com que o agente não reconhecesse a entidade relacionada a intenção de compra. Isso ocorre quando ao treinar o algoritmo de Inteligência Artificial, o

responsável pelo desenvolvimento do sistema não considera a relação de sinônimos do termo, esse processo necessário em algumas plataformas, como o IBM Watson (IBM CLOUD, 2020).

Nesse sentido, ter um tesauro sobre domínio de conhecimento para cujo *chatbot* foi desenvolvido, ajudaria a evitar esse tipo de problema. Na fase de planejamento do agente, e com base nas intenções levantadas, o responsável por desenvolver o diálogo poderia prever quais termos (ou entidades) o *chatbot* assumiria em suas interações e, principalmente, quais os sinônimos ou maneiras o usuário poderia falar a mesma coisa, porém utilizando expressões diferentes. Nesse sentido, a Ciência da Informação, por meio dos Sistemas de Organização do Conhecimento, pode ajudar nesse processo, sendo um deles o tesauro.

Uma das principais dificuldades em organizar o conhecimento, representar e recuperar a informação se deve à problemas relacionados aos fenômenos linguísticos, tais como a sinonímia, a polissemia e a homografia. Sinonímia é quando diferentes termos representam o mesmo conceito. Exemplos de sinonímias são os termos: casa, moraria, lar. Já a polissemia, é o episódio onde um termo possui mais de um significado ou aceitação. Um exemplo disso são os diferentes significados para o termo protocolo que possui significados diferentes na medicina, administração, informática, economia. Por fim, a homografia são termos que possuem a mesma escrita, porém com significados diferentes. Um classifico exemplo é a palavra manga, que pode significar uma fruta, mas também parte de uma camisa. (SCHIESSL; SHINTAKU, 2012).

É possível imaginar o quão complexo é traduzir um documento que está em linguagem natural, para uma linguagem de indexação e que represente por completo determinado conteúdo e, por conseguinte, ajude usuários e sistemas de informação a serem mais eficientes. Nesse contexto, os tesauros têm se mostrado mais eficientes na minimização dos problemas de ambiguidade de uma palavra com significados diferentes, dos significados semelhantes para palavras diferentes, da grafia de uma mesma palavra com diferentes significados.

Schiessl e Shintaku (2012) definem tesauros como:

> Tesauros são vocabulários controlados formados por termos-descritores semanticamente relacionados, e atuam como instrumentos de controle terminológico. Os tesauros podem estar estruturados hierarquicamente (gênero-espécie e todo-parte) e associativamente (aproximação semântica), e são utilizados principalmente para indexar e recuperar informações por meio de seu conteúdo. (DE SALES; CAFÉ, 2009)

Já González (2011) define tesauros como:

> [...] lista de descritores (termos controlados) que representam os conceitos de um domínio do conhecimento. E se organizam em estrutura hierárquica com relações semânticas [...]

Tesauros são, portanto, vocabulários controlados utilizados em um domínio específico e têm como objetivo compilar um documento por meio de termos e, por conseguinte, organizar esses termos por semelhanças (semanticamente) ou por relações lógicas (hierarquização). Os principais objetivos dos tesauros, segundo González (2011) são: eliminar ambiguidades da linguagem em um domínio especifico tal qual pertence os termos; representar de maneira única o conteúdo de um documento; auxiliar o usuário ou o profissional da informação na indexação dos documentos e consultas; controlar o vocabulário utilizado em um determinado sistema documentário; intermediar o uso da informação entre o analista e o usuário.

A estrutura de um tesauro, de modo geral, está dividida em duas partes: base teórica e base técnico-operacional. A base teórica diz respeito a essência e a fundamentação para a construção de um tesauro, incluindo: conceitos, termos, categorias e faceta. Já a base técnico-operacional, aborda as práticas de desenvolvimento dos tesauros, que vai desde o planejamento, coleta de termos, controle terminológico, estabelecimento de relações entre os conceitos e a forma de divulgação e publicação dos mesmos.

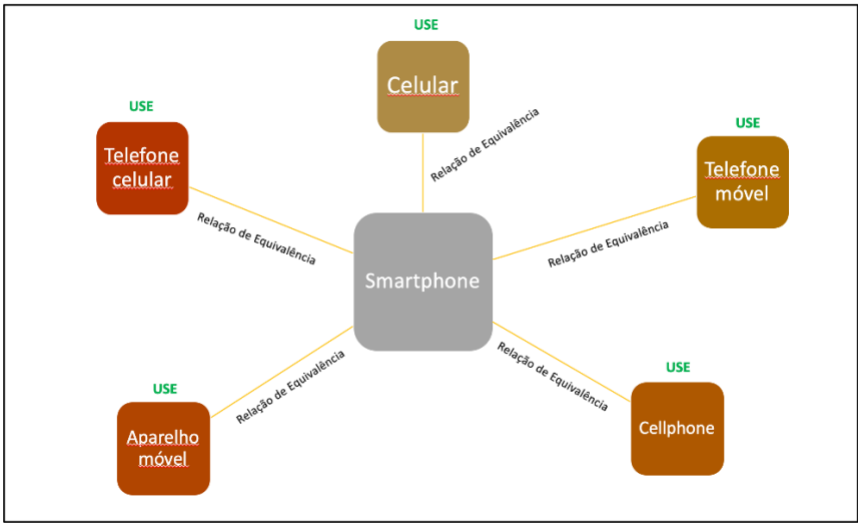
No que tange as relações semânticas dos tesauros, pode-se dizer que existem 3 tipos (CARLAN, 2011; GONZALEZ, 2011): relação de equivalência, relação hierárquica e relação associativa. A relação de equivalência controla os termos utilizados e não utilizados em um tesauro, isto é, estabelece o termo preferido (aquele mais conhecido pelo senso comum) e o menos preferido (aquele termo que é sinônimo do termo preferido, porém não muito utilizado). Já a relação hierárquica apresenta níveis de superordenação (TG – Termo Geral) e subordinação (TE – Termo Específico), algo semelhante a uma estrutura de taxonomia. Por fim, a relação associativa é utilizada quando não se pode identificar uma relação de equivalência ou hierárquica entre os termos.

O tesauro é empregado pelo profissional da informação no momento da representação do documento, isto é, no processo de indexação da informação registrada em linguagem natural para uma linguagem normalizada, pois possibilita a seleção do termo mais adequado, além da sua padronização e a identificação dos relacionamentos entre os termos, o que permite delimitar e tornar a busca mais precisa. Já para o usuário final, mesmo que busque por uma palavra

diferente daquela especificada no momento da representação, a estrutura de relações entre os termos dos tesauros possibilita o "*match*" com o termo semelhante, permitindo assim, a recuperação da informação desejada (SCHIESSL; SHINTAKU, 2012).

Nesse sentido, os tesauros podem ser adotados como metodologia para melhorar o desenvolvimento e a fluidez de uma conversa em um *chatbot*. Um exemplo claro seria utilizar os tesauros para levantar os sinônimos equivalentes para uma ou mais entidades. Ao considerar a entidade "smartphone" como um termo descritor (termo principal), poderia ser listado suas equivalências (representado pela sigla USE), como: celular, telefone celular, telefone móvel, cell, entre outros. Quando o usuário inserir uma pergunta do tipo: "*Como faço para comprar um celular"* ou "*Como faço para comprar um telefone móvel?*", o *chatbot* irá identificar que a intenção para ambas as entradas é "comprar algo" e que as variações entre as entidades "celular" e "telefone móvel" correspondem ao termo descritor "smartphone".

Ilustração 14: Exemplo de tesauro para o termo Smartphone



Fonte: O autor (2020)

Outro benefício ao utilizar os tesauros é no auxílio da identificação de outras intenções e entidades a partir daquelas levantadas no processo de indexação. Por exemplo, considere um *chatbot* especialista em advocacia que será usado tanto por advogados, como por usuários normais. Trata-se de um *chatbot* com um domínio de conhecimento bastante específico e, portanto, seus termos também serão. Com o objetivo do *chatbot* entender tanto o advogado - que provavelmente irá utilizar termos de sua área -, como um usuário comum - que não utilizará termos específicos do setor jurídico -, o profissional responsável por desenvolver esse diálogo

poderá basear-se em tesauros já prontos, como o tesauro jurídico do Supremo Tribunal Federal (STF, 2020). Com esse tesauro, e por meio de suas relações de termos, o profissional poderá ter ideias de novos sinônimos, novas entidades e intenções para serem adicionadas na base de conhecimento do agente, resultando em um diálogo onde será compreendida as perguntas de ambos indivíduos.

Portanto, o tesauro é um método sugestivo para auxiliar na identificação de sinônimos e equivalência de entidades no processo de desenvolvimento de um *chatbot*. Não obstante, os tesauros podem ajudar a agrupar todas as entidades que tenham relação com uma determinada intenção. Ex: **Conta Corrente** e **Conta Poupança** são entidades que poderiam fazer parte da intenção **Consultar Saldo**. Em suma, com essas relações de equivalência, associativas e hierárquicas para uma determinada entidade, o profissional responsável por desenvolver o agente poderá treinar os algoritmos de inteligência artificial do *chatbot* de maneira mais eficiente, de modo que ao interagir com o agente, o sistema reconheça as perguntas dos usuários, mesmo que sejam realizadas de diferentes maneiras.

5.4. Organizando fluxos de diálogo com a Taxonomia

Ao interagir com alguns *chatbots*, é possível observar que muitos não seguem um fluxo de diálogo lógico ou simplesmente perdem-se no meio do caminho conforme as respostas do usuário. Observa-se, por analogia, que esse fenômeno acontece devido à falta de relacionamento hierárquico correto entre os termos (ou entidades e intenções). Uma das etapas importantes no processo de recuperação da informação é justamente esse relacionamento terminológico (ALVARES, 2012). Portanto, após o levantamento das principais intenções e entidades (bem como as suas equivalências e associações), parte-se para uma etapa importante de relacionamento entre eles. A não organização de um fluxo de diálogo coerente, pode levar o usuário a ter uma interação ineficaz, como no exemplo fictício a seguir:

Usuário: Gostaria de pedir uma pizza
Chatbot: Qual o tamanho?
Usuário: Tamanho normal
Chatbot: Ficou R$ 45,00. Algo mais?

Usuário: Calma, eu ainda nem falei o sabor que eu quero, a quantidade, a bebida, meu endereço...

Percebe-se que no diálogo existe uma quebra na sequência da conversa, não existe uma estrutura lógica de perguntas. A resposta apresentada pelo *chatbot* "*Ficou R$ 45,00. Algo mais?*" deveria ser uma das últimas interações, pois as perguntas iniciais como sabor, quantidade e bebida, influenciariam no preço final do pedido. Nota-se a importância de estruturar um diálogo coerente e que siga um fluxo lógico. Essa atividade requer uma certa habilidade em organizar ou relacionar termos (intenções e entidades) de modo que as perguntas e respostas levem o usuário a uma ação assertiva ou à uma informação de qualidade. Nesse sentido, a Ciência da Informação, por meio da taxonomia, pode colaborar para que esse diálogo seja construído de modo a resultar-se em uma a conversa natural com o usuário.

Em seu dia a dia, o homem sempre procura maneiras de organizar o seu pensamento e as coisas que o rodeia. Umas das formas mais naturais de organização do pensamento e das coisas é distinguir e classificar tudo em sua volta por meio das semelhanças e diferenças (SCHIESSL; SHINTAKU, 2012). Essa prática, inerente do ser humano, permite maior compreensão de mundo e convívio em sociedade. Um exemplo é que quando se vai a um supermercado ou uma loja, observa-se que todos os produtos estão organizados em categorias que se assemelham. Tudo isso ajuda para que os consumidores encontrem o que procura de maneira simples. Essa mesma técnica é utilizada no âmbito da informação, já que organizar a informação por semelhança facilita a sua recuperação e visualização. Nesse contexto, uma das principais e mais antigas ferramentas que têm o propósito de organizar a informação, e que também é um Sistema de Organização do Conhecimento, é a taxonomia.

Schiessl e Shintaku (2012) definem taxonomias como:

> Taxonomias são representações semânticas hierarquizadas. Cada classe é uma entidade de informação que está representando o mundo real. Elas são interligadas por um tipo de relação que as une, organiza e também as diferencia. Esse processo de conexão entre as classes estabelece uma hierarquia semântica, pois as entidades estão relacionadas dentro do mesmo contexto semântico.

Já González (2011) a define como:

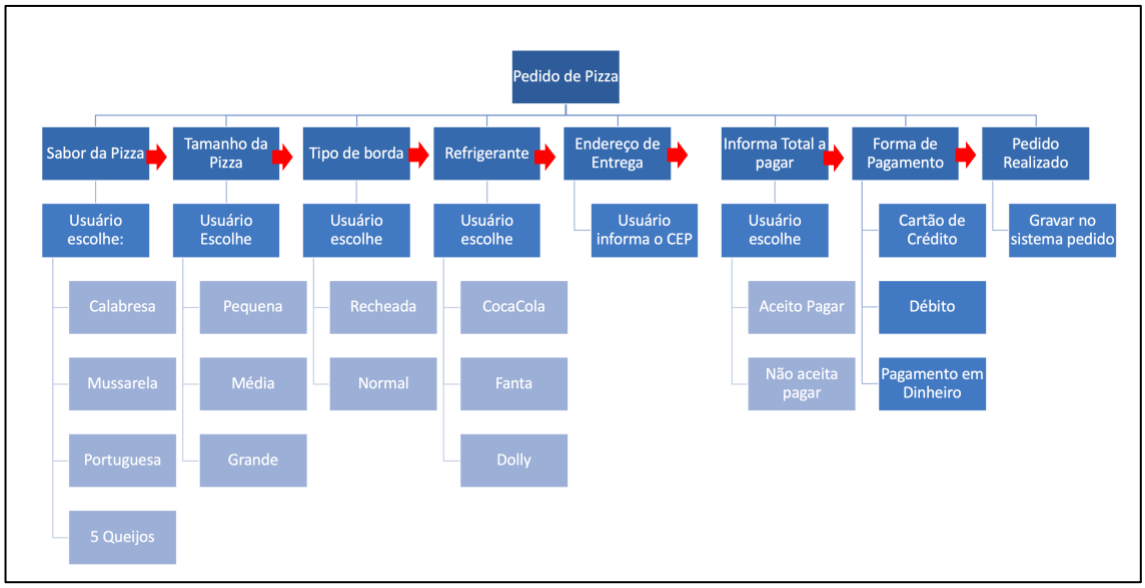
> [...] trata-se de uma lista de termos preferenciais com estrutura hierárquica. Uma taxonomia serve para ordenar informação em uma hierarquia, utilizando a relação pai-filho (generalização ou "tipo de"). Uma hierarquia é basicamente uma estrutura em forma de arvore que estabelece relação entre os objetos de generalização-

> especialização. Uma taxonomia estabelece uma semântica simples de acordo com algumas de suas prioridades.

As taxonomias estão estruturadas em um esquema de árvore hierárquica, ou seja, trata-se de uma estrutura que organiza e agrupa as informações de forma lógica e por tópicos semelhantes (classes e subclasses). Por ter uma estrutura de árvore hierárquica, as informações estruturadas por taxonomias partem de um termo mais abrangente até o mais específico. Observa-se aí um formato no estilo *top-down* (SCHIESSL; SHINTAKU, 2012), indo do termo mais abrangente para o mais específico.

As taxonomias são utilizadas em diversos contextos. Por exemplo, a classificação dos seres vivos (reino, filo, subfilo) é um exemplo claro da aplicação das taxonomias. Além disso, elas também são utilizadas como instrumento de organização e recuperação de informação nas empresas e instituições (GONZALEZ, 2011). As bibliotecas, ao agruparem livros por área do conhecimento e, por conseguinte, subáreas desse conhecimento, também utilizam as taxonomias. Páginas da internet que possuem menus hierárquicos é um outro exemplo de uso de taxonomias. Para o desenvolvimento de um *chatbot*, a utilização de uma taxonomia pode ser algo bastante eficiente para que o profissional possa visualizar e construir fluxos de diálogos mais coerentes.

Ilustração 15: Taxonomia para um *Chatbot* de pedido de pizza



Fonte: O autor (2020)

A Ilustração 15 mostra um exemplo bem simples da utilização da taxonomia para construir e visualizar como seria uma estrutura de diálogo para um pedido de pizza. Veja que para finalizar

a compra, o *chatbot* precisa pedir diversas informações para o usuário. Conforme o indivíduo responde à essas perguntas, o fluxo vai avançando até que o sistema colete todos os dados, confirme-os com o usuário e, finalmente, por meio de uma API, o sistema de *chatbot* possa transferir esses dados para um sistema de solicitação de pizza da empresa.

Como se observa, a taxonomia pode ser um modelo de visualização de informação essencial no processo organizacional de intenções e entidades, já que ao organizar as informações de forma lógica e hierárquica, a compreensão sobre o todo é facilitada, permitindo uma melhor visualização de como o usuário poderia interagir em uma conversa com um *chatbot*. Não obstante, essa taxonomia poderá ser compartilhada entre todos os envolvidos no processo de desenvolvimento do agente. Com uma estrutura clara do diálogo, todos poderão sugerir melhorias ou a adição de um novo nível na hierarquia da conversa. Desta forma, o desenvolvedor terá uma estrutura mais próxima do ideal para, finalmente, desenhá-la dentro de uma plataforma de desenvolvimento.

5.5. Utilizando a Ontologia no relacionamento de conversas

Uma das principais razões que leva o *chatbot* a ter uma interação satisfatória com o usuário é sua capacidade de manter uma conversa natural (ARTIFICIAL SOLUTIONS, 2020). Isso corresponde a manter um diálogo que supra as expectativas do usuário e, principalmente, o auxilie com informações extras que levarão a uma recuperação da informação ainda mais eficiente. Considere o exemplo:

Usuário: Olá, gostaria de saber a previsão de tempo para amanhã.
Chatbot: Claro, onde você está localizado?
Usuário: Sou de São Paulo capital.
Chatbot: Amanhã vai fazer sol em São Paulo. Aqui vai uma dica: visitar a USP na Cidade Universitária é uma ótima opção para aproveitar o solzinho. Quer dicas de como chegar lá?
Usuário: Nossa, como você é proativo. Quero sim!

Observe que o diálogo é um claro exemplo hipotético de como um *chatbot* pode agregar valor no momento que o usuário interage com ele. Veja que, ao perceber que fará sol em São Paulo, o agente recomenda que o usuário visite a Cidade Universitária da USP como uma opção de lazer para aproveitar a temperatura. Esse tipo de sugestão pode ser planejado no momento de

desenvolver a estrutura de um diálogo. Nesse sentido, a Ciência da Informação e a Ciência da Computação podem colaborar nesse processo com o uso das Ontologias que, sob o ponto de vista da CI, é considerado um dos Sistemas de Organização do Conhecimento.

A Ciência da Informação é uma ciência interdisciplinar que se relaciona com várias áreas do conhecimento. Isso se dá pelo fato do seu objeto – informação – ser bastante abrangente e, muitas vezes, empregado como tema de pesquisa em alguns estudos ou áreas. Nesse contexto, a Ciência da Computação e as Engenharias possuem laços fortes com a CI, sobretudo a área de organização e recuperação da informação. Algumas técnicas de representação do conhecimento são compartilhadas entre elas, sendo uma delas, a ontologia.

Uma das definições mais disseminadas sobre ontologia é a de Gruber (1993, p1) que a define como: "uma especificação explícita de uma conceptualização". Nesse sentido, a palavra "explícita" se deve ao fato que os objetos simbólicos devem ser definidos de maneira formal e clara. Já a palavra "conceptualização" diz respeito a composição de objetos, seus conceitos e suas relações em torno de um determinado domínio (BOCCATO; RAMALHO, 2008). Schiessl e Shintaku (2012) resumem a ontologia da seguinte forma:

> [...] pode-se entender a ontologia como o meio de representação de determinada realidade pela conceituação sistemática, compartilhada e formal, dentro de um domínio do conhecimento, mas sem restrições disciplinares, buscando a padronização dos termos e conceitos. Assim, tem-se uma visão ampla da ontologia, não propriamente como um artefato, mas como uma base conceitual.

Já Ramalho (2006) define ontologia da seguinte maneira:

> Um artefato tecnológico que descreve um modelo conceitual de um determinado domínio em uma linguagem lógica e formal, a partir da descrição dos aspectos semânticos de conteúdos informacionais, possibilitando a realização de inferências automáticas por programas computacionais.

González (2011) define ontologia como:

> Ontologia é um conceito filosófico que estuda o ser em geral, suas propriedades transcendentes e, mais especificamente, a organização da realidade. Na representação da informação, pode entender-se como a soma de uma série de conceitos relevantes que representam o conhecimento compartilhado pelos membros de um determinado domínio.

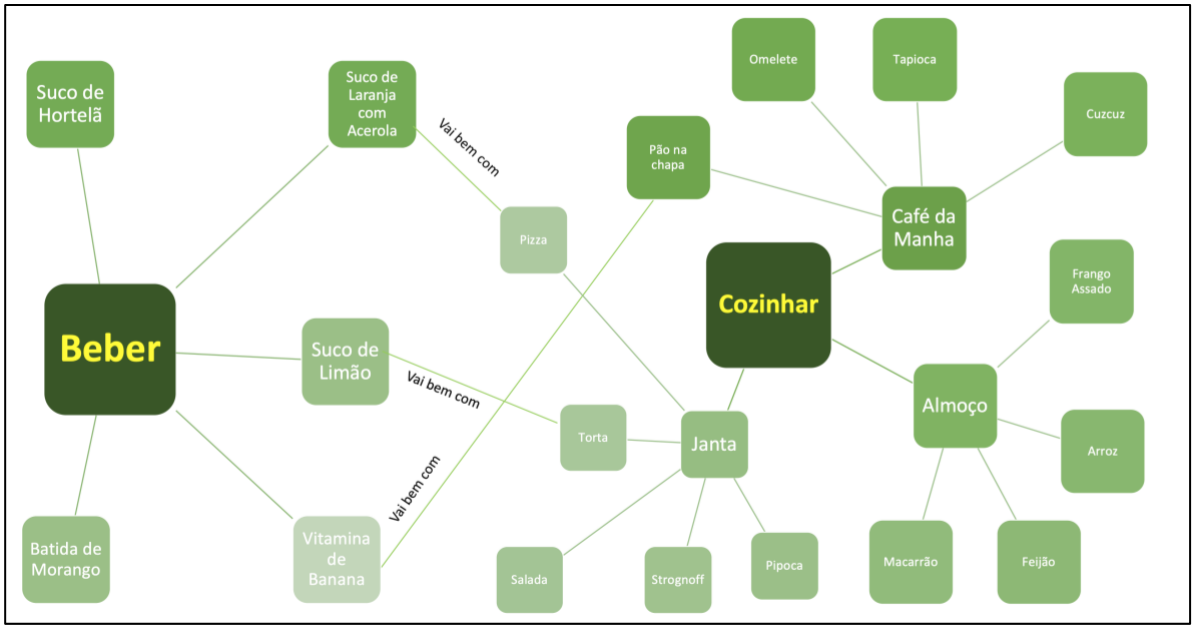
De forma sucinta, a ontologia é uma ferramenta para a representação e recuperação da informação, está relacionada com a recuperação de forma automatizada (informática). Nesse contexto, a ontologia tem uma relação bastante intrínseca com o tesauro, essencialmente pela

sua estrutura de relacionamento entre os termos. A diferença, entretanto, é que o tesauro nasce no domínio da Ciência da Informação, enquanto a ontologia nasce no seio da Ciência da Computação, especificamente no contexto da Inteligência Artificial, pela necessidade de descrever e representar a visão de mundo das entidades e suas relações computacionais (SCHIESSL; SHINTAKU, 2012).

Essa explanação de origens, causa impacto em algumas diferenças práticas entre as duas técnicas. A preocupação primordial do tesauro está na definição de termos como um instrumento para registro terminológico e representação da informação, sendo utilizado essencialmente pelos seres humanos. Já a ontologia se pauta muito na parte de formalização, algo essencial para as inferências em processamento de sistemas automatizados (SCHIESSL; SHINTAKU, 2012).

Ramalho (2006) destaca que as ontologias possuem estruturas mais complexas de relações, o que possibilita descrever níveis de relacionamentos mais avançados e que podem ser processados de forma automatizada por computadores, sobretudo pela preocupação na formalização dos objetos. Nesse contexto, hoje, as ontologias são utilizadas amplamente em sistemas de recuperação de informação, especialmente aqueles que utilizam arquiteturas de inteligência artificial, já que possibilitam conexões de forma mais ampla, além de viabilizar a interoperabilidade entre o banco de dados. Veja a Ilustração 16 que mostra como uma ontologia poderia ser aplicada no planejamento de relacionamento entre intenções e entidades de um *chatbot*.

Ilustração 16: Ontologia para um *Chatbot* de Receitas



Fonte: O autor (2020)– Adaptação Masycheff (2020)

Observe que a Ilustração 16 trata de um diálogo desenvolvido em um contexto em que um *chatbot* fornece informações sobre receitas. Dentro da intenção **cozinhar** existem as entidades **café da manhã**, **almoço** e **janta**. Cada uma das entidades possui entidades de segundo nível que representam o tipo de prato para cada ocasião. Assim, ao perguntar algo como: "Quero fazer uma Pizza", o agente irá identificar a intenção de **cozinhar** e, automaticamente, por meio das relações, também identificará as entidades relacionadas a **jantar** e a **pizza**. Além disso, graças as relações proporcionadas pela ontologia, o *chatbot* poderá considerar que pizza é uma boa opção com suco de laranja com acerola. Ao identificar essa relação, o agente pode retornar a seguinte pergunta ao usuário: "*Quer aproveitar e aprender como fazer um delicioso suco de laranja com* acerola? Normalmente é uma boa pedida com pizza". Ao responder de forma positiva, o sistema fornecerá ao indivíduo uma receita de pizza e, adicionalmente, uma receita de suco de laranja com acerola.

Além de auxiliar no processo de construção do fluxo de diálogo, a ontologia pode ajudar o profissional com a visualização do processo de interoperabilidade entre o sistema de conversa do *chatbot* e um sistema externo, podendo ser de terceiros ou da própria empresa. Considere, por exemplo, um agente de uma empresa de reserva de passagens aéreas. Trata-se de um *chatbot* que necessitará obter informações de diversos sistemas distintos, tais como: sistema de reserva, sistema de pagamento, sistema de hotéis, etc. Por meio de uma ontologia, o profissional poderá

ter ideais mais claras sobre em qual momento do diálogo ou tipo de pergunta do usuário, o *chatbot* requisitará (por meio de uma API) um outro sistema.

Em suma, a ontologia é o tipo de Sistema de Organização do Conhecimento mais complexo e completo que existe (ALVARES, 2012). Por meio dela, pode-se criar diversos tipos de relações que corroborará na visualização que o diálogo irá transcorrer entre o usuário e o *chatbot*. Por ser um modelo de organização do conhecimento pronto para ser interpretado por máquinas, um desenvolvedor poderá construir todo o diálogo através de um sistema de construção de ontologia, como o *Protégé*, e conectá-lo direto com um *chatbot*. Ou seja, com o uso de uma API que se conecta a um servidor de ontologia, o agente recebe a pergunta do usuário, identifica a intenção, consulta essa intenção na ontologia e, logo em seguida, a ontologia volta com um valor que é repassado para o usuário em forma de resposta (SILVA, 2019). Portanto, além de um modelo conceitual para apoiar na visualização completa de um fluxo diálogo, esse mesmo modelo pode ser aproveitado diretamente nas plataformas de desenvolvimento de *chatbot*.

6. CONSIDERAÇÕES FINAIS

Utilizando-se de uma linguagem acessível para aproximar a academia e o mercado profissional, a presente pesquisa teve por objetivo evidenciar como os *chatbots* colaboram para a efetiva recuperação da informação, processo estudado pela Ciência da Informação, o que levou-se a procurar identificar, em contrapartida, como a Ciência da Informação colabora para o desenvolvimento e aprimoramento desses sistemas. Iniciou-se a pesquisa explorando o campo da Ciência da Informação e suas principais vertentes. Versou-se sobre o que são os *chatbots* e suas principais tecnologias e tendências. Na sequência, analisou-se alguns casos de uso desses agentes sob o ponto de vista de sua colaboração em facilitar a recuperação da informação. Por fim, o último capítulo percorreu e evidenciou como algumas ferramentas da CI contribuem para o desenvolvimento e aprimoramento desses sistemas.

Conforme explorado nesta pesquisa, o uso dos *chatbots* configuram em uma sugestão de ferramenta para facilitar a recuperação da informação. Por meio da Inteligência Artificial, sobretudo com uso da Linguagem Natural, esses agentes permitem que os usuários possam interagir e satisfazer suas necessidades de informação de maneira simples e menos burocrática. Embora muitas informações fornecidas por esses agentes possam ser encontradas nos websites dessas instituições, a obtenção delas é facilitada por meio de uma simulação de conversa. Isso explica o porquê os *chatbots* mostram-se como uma tendência de uso e aplicação em diversos setores, seja auxiliando as pessoas em um processo de compra, ajudando com a identificação de doenças ou interagindo com elas em um museu de arte.

Como aponta o HypeCycle, os *chatbots* estão na fase de mostrar a sua real utilidade, ou seja, essa é a etapa na qual as empresas estão testando a sua eficácia. Conforme os resultados dessas aplicações provam-se satisfatórios, o uso desses agentes em sociedade se tornará cada vez mais comum. Dessa maneira, cada vez mais as empresas os utilizarão em seus processos, sobretudo com a finalidade de automatizar o atendimento ao cliente. Esses agentes serão cada vez mais comuns nas residências das pessoas, comportando-se como assistentes virtuais. Esse movimento já pode ser visto atualmente com a popularização de aparelhos residências como o da Apple (Siri), Google (Home) e Amazon (Echo). Entretanto, para que esses agentes permaneçam em evidência, eles precisam oferecer uma boa experiência para os usuários, compreender e satisfazer as suas necessidades de informação. Nesse ínterim, a Ciência da Informação passa a ser uma importante área do conhecimento para corroborar esse processo.

Conforme visto, um dos principais desafios encontrado em um *chatbot* é justamente no momento que o usuário se comunica com eles. Muitos desses agentes tendem a apresentar problemas de interação, o que inclui questões como: a não compreensão da pergunta inserida pelo usuário; tom de conversa em desconexão com as características do indivíduo; diálogos confusos e informações inadequadas; a não distinção entre palavras com mesmos sinónimos, entre outros. Tal qual versado nesta pesquisa, muito desses problemas podem estar relacionados a assuntos como: a carência em compreender o usuário; a falta de um trabalho inicial de indexação de conteúdo; o uso de controle de vocabulário; a coleta e relacionamento entre termos. Todos esses assuntos são estudados há bastante tempo na Ciência da Informação. Logo, ao incorporar os instrumentos e processos da CI ela contribui para diminuir esses problema.

Como se observa nas passagens de Seth Ealey no decorrer deste texto, não existe inteligência artificial sem arquitetura de informação; isto porque os sistemas de Inteligência Artificial, incluindo aqui os *chatbots*, necessitam ser alimentados com informações organizadas para que, de fato, tenham um comportamento inteligível. Para que esses algoritmos desempenhem um papel satisfatório e, por conseguinte, compreenda e satisfaça as necessidades de informações do usuário, é necessário, inicialmente, um importante trabalho humano. Esse trabalho inclui, primeiramente, compreender as necessidades do usuário para que haja um trabalho de coleta, organização e disponibilização dessas informações. Portanto, os subdomínios da CI, tais como estudos de usuário da informação, indexação, taxonomias, tesauros e ontologias, configuram-se em técnicas e ferramentas que podem auxiliar na melhoria e desenvolvimento dos *chatbots*. Nesse sentido, o profissional da informação também se torna um agente importante nesse processo.

O profissional da informação, como o bibliotecário e o indexador, é o sujeito que possui conhecimento teórico e prático sobre o fenômeno informação. Durante a sua trajetória acadêmica, esse indivíduo aprende a refletir sobre o uso e disseminação da informação e, principalmente, a compreender o usuário que irá consumi-la. Ter esse tipo de profissional como membro de uma equipe de desenvolvimento de *chatbots*, parece ser algo bastante sugestivo. Ao possuir esse arcabouço de conhecimento, esse sujeito pode colaborar de forma contundente no processo de curadoria e organização da informação; processo esse que, como evidenciado, é muito importante para o bom desempenho de um sistema de inteligência artificial, como um

chatbot. Posto que, são esses os profissionais que tendem a dominar as técnicas da CI que foram apresentadas aqui como sugestões para melhorar a inteligência desses agentes.

Enfim, tentou-se evidenciar aqui diversos aspectos de como os *chatbots* podem ser uma ferramenta útil para a recuperação da informação e, principalmente, como a Ciência da Informação pode empregar um papel importante no processo de desenvolvimento desses agentes. Entretanto, se faz necessário mencionar que esta pesquisa não teve como intuito esgotar os assuntos levantados, mas sim que as ideias discutidas aqui reverberem para que novos estudos surjam. Testar a eficiência das ferramentas de CI apresentadas como sugestão no processo de desenvolvimento de um *chatbot* poderia ser um tipo de estudo futuro que continuaria este trabalho. Além disso, e com base nas normativas da área - tais como a ANSI/NISO Z39.19:2005 que se refere a construção de ontologias e tesauros, bem como a ABNT NBR, 12676 que auxilia no processo de indexação – é possível arquitetar um manual prático de como melhorar a base de conhecimento de um *chatbot* com o uso de tesauros, ontologias e taxonomias.

Em síntese, os *chatbots*, sobretudo aqueles que utilizam elementos da inteligência artificial, mostram-se como uma importante tendência para diversas empresas e setores. Existem muitas plataformas de desenvolvimento com o objetivo de facilitar a criação desses agentes. Muitas delas não requerem que o profissional tenha sequer conhecimento em programação. Entretanto, compreender o fenômeno informação deveria ser algo essencial nesse processo. Isto é, o cerne dessas novas tecnologias é a informação. Entender todos os aspectos desse fenômeno, desde a sua coleta e organização até os questionamentos mais profundos de seu uso e impactos culturais, seria algo essencial para o progresso dessas tecnologias. Ferramentas tecnológicas surgem a cada instante, porém se faz necessário ter inteligência informacional humana para torná-las artificialmente inteligentes a favor da sociedade. Sem isso, as promessas e os benefícios do *chatbot* e da inteligência artificial configuram-se apenas em promessas.

REFERÊNCIAS

A VOZ DA ARTE - YOUTUBE. **A Voz da Arte**. Disponível em: https://www.youtube.com/watch?v=1rOAgvCnZpw. Acesso em: 04 ago. 2020.

ALENCAR, Antônio Juarez; SCHMITZ, Eber Assis; CRUZ, Leôncio Teixeira. **Assistentes Virtuais Inteligentes e Chatbots:** Um guia prático e teórico sobre como criar experiências e recordações encantadoras para os clientes da sua empresa. Brasport, 2018.

ALMEIDA, Daniela Pereira dos Reis de et al. Paradigmas contemporâneos da Ciência da Informação: a recuperação da informação como ponto focal. Revista Eletrônica Informação e Cognição, v. 6, n. 1, p. 16-27, 2007.

Alvares (ORG.). Lillian. **Organização da Informação e do Conhecimento**: Conceitos, subsídios, interdisciplinaridades e aplicações. São Paulo: B4editores, 2012.

ALVES, Rachel C. V. et al. **Ciência da Informação, Ciência da Computação e Recuperação da Informação**: algumas considerações sobre os métodos e tecnologias da informação utilizados ao longo do tempo. Revista Eletrônica Informação e Cognição, Marília, SP, v.6, n.1, p.28-40, 2007.

ALVES, Roberta Caroline Vesú et al. **Ciência da Informação e a Pós-Modernidade**: considerações sobre o status científico. Revista Eletrônica Informação e Cognição, v. 6, n. 1, p. 41-54, 2007. Acesso em: 21 apr. 2019

AMAZON. **Alexa Skills Kit Official Site**: build skills for voice. Build Skills for Voice. 2020. Disponível em: https://developer.amazon.com/pt-BR/alexa/alexa-skills-kit. Acesso em: 29 jul. 2020.

ANSI/NISO **Z39.19-2005**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO, 2005. 172 p. Disponível em: <http://www.niso.org/kst/reports/standards/kfile_download?id%3Austring%3Aiso-8859-1=Z39-19-2005.pdf&pt=RkGKiXzW643Y eUaYUqZ1BFwDhIG4-24RJbcZBWg8uE4vWdpZsJDs4RjLz0t90_d5_ymGsj_IKVaGZww13HuDlYn5U74YdfA-3TffjxYQ25QrtR8PONuJLqxvo-l0NIr5>. Acesso em: 10 jul. 2020.

ARAÚJO, Carlos Alberto Ávila. **Paradigma social nos estudos de usuários da informação**: abordagem interacionista. 2012.

ARTIFICIAL SOLUTIONS. **Chatbots**: The Definitive Guide (2020). Disponível em: https://www.artificial-solutions.com/chatbots. Acesso em: 10 maio 2020.

ATAÍDE DIAS, Guilherme; HENN, Gustavo; WENDELL DE MORAIS SILVA, José. **Tecnologia da informação e serviços de referência eletrônicos**: uma proposta de aplicação baseada em chatbots e ontologias. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, n. 23, 2007.

ATENDIMENTO MAGALU. **Atendimento Magalu**: a voz da arte. A Voz da Arte. Disponível em: https://atendimento.magazineluiza.com.br/hc/pt-br. Acesso em: 04 ago. 2020.

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Recuperação de Informação**: Conceitos e Tecnologia das Máquinas de Busca. Bookman Editora, 2013.

BAPTISTA, Sofia Galvão; CUNHA, Murilo Bastos da. **Estudo de usuários**: visão global dos métodos de coleta de dados. Perspectivas em ciência da informação, v. 12, n. 2, p. 168-184, 2007.

BARRETO, Aldo de Albuquerque. **Uma história da ciência da informação**. Para entender a ciência da informação. Salvador: EDUFBA, p. 13-34, 2007.

BARTH, Fabricio. **Alguns detalhes técnicos e desafios do projeto "A voz da arte – IBM | Watson". 2017**. Disponível em: https://www.linkedin.com/pulse/alguns-detalhes-técnicos-e-desafios-do-projeto-voz-da-fabr%C3%ADcio-barth/. Acesso em: 30 maio 2020.

BAX. M.; DIAS, E. **A abordagem da "construção de sentido" em estudos de usuários**. Belo Horizonte: Escola de biblioteconomia da UFMG; NAPq, Núcleo de Assessoramento à Pesquisa, 1997. Cadernos de Ciência da Informação, v.1.

BOCCATO, Vera Regina Casari; RAMALHO, Rogério Aparecido Sá; FUJITA, Mariângela Spotti Lopes. **A contribuição dos tesauros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais**. Ibersid: revista de sistemas de información y documentación, v. 2, p. 199-209, 2008.

BORKO, H. **Information Science**: What is it? American Documentation, v.19, n.1, p.3-5, jan. 1968.

BOTS BRASIL. **Meu robô não se confunde**. Disponível em: https://www.google.com/search?client=safari&rls=en&q=meu+robo+nao+se+confunde&ie=UTF-8&oe=UTF-8. Acesso em: 10 jul. 2020.

BUCKLAND, Michael K. **Information as thing**. Journal of the American Society for information science, v. 42, n. 5, p. 351-360, 1991.

BUCKLAND, Michael K. **What is a "document"?** Journal of the American society for information science, v. 48, n. 9, p. 804-809, 1997.

BUCKLAND, Michael. **What kind of Science can Information Science be?** Journal of the American Society for Information Science and Technology (ASIS&T), v, 63, n.1, p. 1-7, 2012. Disponível em: Acesso em: https://onlinelibrary.wiley.com/doi/epdf/10.1002/asi.21656 .10 jun. 2019.

BURCH, Sally. **Sociedade da informação/sociedade do conhecimento**. Ambrosi, A.; Peugeot, V.; Pimenta, D. Desafios das palavras. Ed. VECAM, 2005.

BURKE, Peter. **Uma história social do conhecimento 1**: de Gutenberg a Diderot. Zahar, 2003.

BURKE, Peter. **Uma história social do conhecimento 2**: da Enciclopédia à Wikipédia. Zahar, 2012.

CAFEZEIRO, Isabel; COSTA, Leonardo Cruz da; KUBRUSLY, Ricardo da Silva. **Ciência da Computação, Ciência da Informação, Sistemas de Informação**: Uma reflexão sobre o papel da informação e da interdisciplinaridade na configuração das tecnologias e das ciências. Perspectivas em Ciência da Informação, v. 21, n. 3, p. 111-133, 2016.

CALADO, Caio. **Magazine Luiza — entrevista com o time responsável pela criação da Lu.**: por que testar um chatbot como novo canal de atendimento? quais fatores são fundamentais para se criar uma boa experiência?. Por que testar um chatbot como novo canal de atendimento? Quais fatores são fundamentais para se criar uma boa experiência?. 2018. Disponível em: https://medium.com/botsbrasil/magazine-luiza-entrevista-com-o-time-responsável-pela-criação-da-lu-8fc987fbafad. Acesso em: 31 maio 2020.

CALADO, Caio. **Quer conhecer ou testar os bots finalistas do Bots Brasil Awards 18–19?**: junto a toda a comunidade, vamos votar nos melhores bots do ano que foram nomeados pelo time de especialistas. Junto a toda a comunidade, vamos votar nos melhores bots do ano que foram nomeados pelo time de especialistas. 2019. Disponível em: https://medium.com/botsbrasil/quer-conhecer-ou-testar-os-bots-finalistas-do-bots-brasil-awards-18-19-2ae97cf13523. Acesso em: 01 jun. 2020.

CAPURRO, Rafael; HJORLAND, Birger. **The concept of information as we use in everyday**. Perspectivas em ciência da informação, v. 12, n. 1, p. 148-207, 2007.

CAPURRO, Rafael. **Epistemologia e Ciência da Informação**. In: Encontro Nacional de Pesquisa em Ciência da Informação – ENANCIB, 5. Anais... Belo Horizonte: ECI/UFMG, 2003. Disponível em: http://www.capurro.de/enancib_p.htm. Acesso em: 08 mar. 2019.

CARDOSO FILHO, Jair Cunha; SANTOS, Márcia Mazo. **Processos e temas selecionados. Organização da informação e do conhecimento**: conceitos, subsídios interdisciplinares e aplicações. São Paulo: B4 Editores, p. 185-223, 2012

CARLAN, Eliana; MEDEIROS, Marisa Bräscher Basílio. **Sistemas de Organização do Conhecimento na visão da Ciência da Informação**. 2011.

CASATTI, Denise. **CheckCorona**: esforço de pesquisador da USP é reconhecido por plataforma da ONU: robô que identifica sintomas da covid-19 via WhatsApp foi selecionado para integrar iniciativa que mapeia soluções inovadoras de resposta à pandemia. Robô que identifica sintomas da Covid-19 via WhatsApp foi selecionado para integrar iniciativa que mapeia soluções inovadoras de resposta à pandemia. 2020. Disponível em: https://www.icmc.usp.br/noticias/4929-checkcorona-esforco-de-pesquisador-da-usp-e-reconhecido-por-plataforma-da-onu. Acesso em: 20 maio 2020.

CHATFUEL. **Chatfuel Documentation**. Disponível em: https://docs.chatfuel.com/en/. Acesso em: 11 maio 2020.

CHIOVATTO, Milene. **Watson, uso de Inteligência Artificial (AI) e processos educativos em museus**. Revista Docência e Cibercultura, v. 3, n. 2, p. 217-230, 2019.

CINTRA, Anna Maria Marques et al. **Para entender as linguagens documentárias**. São Paulo: Polis, 2002.

COELHO, Elisa Quaresma; COELHO, Augusto Quaresma; CARDOSO, José Eduardo Dias. **Informações médicas na internet afetam a relação médico-paciente?**. Revista Bioética, v. 21, n. 1, p. 142-149, 2013.

COMARELLA, R. Café, L. **Chatbot**: conceito, características, tipologia e construção, Informação e Sociedade, vol. 18, n.2, 2008.

COMPUTERWORLD. **BIA, do Bradesco, já soma 100 milhões de interações**. 2019. Disponível em: https://computerworld.com.br/2019/03/22/bia-do-bradesco-ja-soma100-milhoes-de-interacoes/. Acesso em: 10 mar. 2020.

DA CUNHA, Murilo Bastos. **Metodologias para estudo dos usuários de informação científica e tecnológica**. Revista de Biblioteconomia de Brasília, v. 10, n. 2, 1982.
da informação. Ciência da Informação, v. 26, n. 1, 1997.

DALE, Robert. **The return of the chatbots**. Natural Language Engineering, v. 22, n. 5, p. 811-817, 2016.

DAY, Ron. **LIS, method, and postmodern science**. Journal of Education for Library and Information Science, v. 37, n. 4, p. 317-324, 1996. Disponível em: https://www.jstor.org/stable/pdf/40324240.pdf. Acesso em: 14 fev. 2019.

DE CANEDA QUEIROZ, Daniela Gralha; DE MOURA, Ana Maria Mielniczuk. **Ciência da Informação**: história, conceitos e características. Em Questão, v. 21, n. 3, p. 26-42, 2015. Disponível em: https://seer.ufrgs.br/EmQuestao/article/viewFile/57516/36041. Acesso em Acesso em: 18 fev. 2019.

DE OLIVEIRA, Hilário TA et al. **Dr. Pierre**: Um Chatbot com Intenção e Personalidade Baseado em Ontologias para Apoiar o Ensino de Psiquiatria. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2010.

DE PONTES JUNIOR, João; DE AQUINO CARVALHO, Rodrigo; AZEVEDO, Alexander Wilian. **Da recuperação da informação à recuperação do conhecimento**: reflexões e propostas. Perspectivas em Ciência da Informação, v. 18, n. 4, p. 02-17, 2013.

DE SALES, Rodrigo; CAFÉ, Ligia. **Semelhanças e diferenças entre tesauros e ontologias**. 2008.

DERVIN, B.;**An overview of sense-making research**: concepts, methods and results to date. In: International Communications Association Annual Metting. Dallas, Texas, 1983.

DIAS, Maria Matilde Kronka. **Usos e usuários da informação**. EdUFSCar, 2004.

EARLEY, Seth. **There is no AI without IA**. IT Professional, v. 18, n. 3, p. 58-64, 2016.

EXAME. **Inteligência artificial faz obra de arte conversar com público**. 2017. Disponível em: https://exame.com/tecnologia/inteligencia-artificial-da-ibm-faz-obra-de-arte-conversar-com-publico/. Acesso em: 31 maio 2020.

EXCEED.AI. **The Benefits of Using an Email Bot Alongside a Chatbot**. Disponível em: https://exceed.ai/email-bot-alongside-chatbot/. Acesso em: 11 maio 2020.

FERNEDA, Edberto. **Recuperação da informação**: análise sobre a contribuição da Ciência da computação para a Ciência da Informação. Tese (doutorado em comunicação) – USP. Escola de Comunicação e Artes, São Paulo, 2003.

FIGUEIREDO, Nice Menezes de. **Estudos de uso e usuários da informação**. 1994.

FREIRE, Gustavo Henrique. **Ciência da informação**: temática, histórias e fundamentos. Perspectivas em ciência da informação, v. 11, n. 1, p. 6-19, 2006.

FUJITA, Mariângela Spotti Lopes. **A identificação de conceitos no processo de análise de assunto para indexação**. Revista digital de biblioteconomia e ciência da informação, v. 1, n. 1, 2003.

GABRIEL, Martha. **Você, eu e os robôs**: pequeno manual do mundo digital. **São Paulo: Atlas**, 2018.

GALVAO, Adjamir M. et al. **Persona-AIML**: an architecture for developing chatterbots with personality. In: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, 2004. AAMAS 2004. IEEE, 2004. p. 1266-1267.

GARTNER. **Gartner Glossary**. Disponível em: https://www.gartner.com/en/information-technology/glossary/chatbot. Acesso em: 11 maio 2020.

GARTNER. **Gartner Hype Cycle**. Disponível em: https://www.gartner.com/smarterwithgartner/top-trends-on-the-gartner-hype-cycle-for-artificial-intelligence-2019/. Acesso em: 10 maio 2020.

GARTNER. **Market Guide for Virtual Customer Assistants**. 2019. Disponível em: https://www.gartner.com/doc/reprints?ct=190715&id=1-1OA3N7O0&st=sb. Acesso em: 10 maio 2020.

GARTNER. **Top Trends on the Gartner Hype Cycle for Artificial Intelligence**, 2019. 2019. Disponível em: https://www.gartner.com/smarterwithgartner/top-trends-on-the-gartner-hype-cycle-for-artificial-intelligence-2019/. Acesso em: 10 maio 2020.

GIL, Antonio Carlos. **Metodologia científica**. São Paulo, v. 3, 2002.

GOL LINHAS AÉREAS. **Site da Gol Linhas Aéreas**. Disponível em: http://www.voegol.com.br. Acesso em: 04 ago. 2020.

GOMES, Murilo S. et al. **Uma Ontologia de Domínio no Contexto de Observatórios**. In: ONTOBRAS. 2016. p. 204-209.

GONZÁLEZ, José Antonio Moreiro. **Linguagens documentárias e vocabulários semânticos para a web**: elementos conceituais. SciELO-EDUFBA, 2011.

GOOGLE CLOUD. **DialogFlow Documentation**. Disponível em: https://cloud.google.com/dialogflow/docs. Acesso em: 10 maio 2020.

GRUBER, Thomas R. **Toward principles for the design of ontologies used for knowledge sharing?**. International journal of human-computer studies, v. 43, n. 5-6, p. 907-928, 1995.

HENN, Gustavo. **Chatbots no serviço de referência online**: uma ferramenta para a gestão da biblioteca da PRT 13ª Região. 2006.

IBM CLOUD. **Criando Entidades**. Disponível em: https://cloud.ibm.com/docs/assistant?topic=assistant-entities&locale=pt-BR. Acesso em: 10 jul. 2020.

IBM CLOUD. **Getting started with Watson Assistant**. Disponível em: https://cloud.ibm.com/docs/assistant?topic=assistant-getting-started. Acesso em: 10 maio 2020.

IBM. **Com BIA, Bradesco e IBM transformam o atendimento de milhões de usuários**. 2019. Disponível em: https://www.ibm.com/blogs/ibm-comunica/com-bia-bradesco-e-ibm-transformam-o-atendimento-de-milhoes-de-usuarios/. Acesso em: 31 maio 2020.

IBM. **Construção do manual cognitivo do Virtus**, Volkswagen. Disponível em: https://mediacenter.ibm.com/media/1_8ak7c7gt. Acesso em: 31 maio 2020.

IBM. **Speech Recognition**: Enable systems and applications to understand and process human speech. Disponível em: https://www.ibm.com/topics/speech-recognition. Acesso em: 10 maio 2020.

KIELGAST, Soeren; HUBBARD, Bruce A. **Valor agregado à informação**: da teoria à prática. Ciência da informação, v. 26, n. 3, 1997.

KOTLER, Philip; KARTAJAYA, Hermawan; SETIAWAN, Iwan. **Marketing 4.0**: do tradicional ao digital. Sextante, 2017.

KUHN, T. S. **Estrutura das revoluções científicas**. 7. ed.. São Paulo: Perspectiva, 2003.

LAVEN, S. **What is a chatterbot?** The Simon Laven page. Disponível em <http://www.simonlaven.com >. Acesso em: 20 maio 2020.

LEONHARDT, M. D. **Um estudo sobre Chatterbots**. 2005. Trabalho individual – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

LEONHARDT, Michelle Denise. **Doroty**: um chatterbot para treinamento de profissionais atuantes no gerenciamento de redes de computadores. 2005.

LIDDY, Elizabeth D. **Natural language processing**. 2001.

LIMA, Gercina Ângela Borém. **Interfaces entre a ciência da informação e a ciência cognitiva**. Ciência da Informação, v. 32, n. 1, 2003.

MANUAL COGNITIVO - YOUTUBE. **Manual Cognitivo**. Disponível em: https://www.youtube.com/watch?v=bTnZFEtSKSw. Acesso em: 04 ago. 2020.

MANCINI, M. C. Estudos de Revisão Sistemática: um guia para síntese criteriosa da evidência científica. 2007. 2017.

MARCONDES, Carlos Henrique; DE ALMEIDA CAMPOS, Maria Luiza. **Ontologia e web semântica**: o espaço da pesquisa em ciência da informação. PontodeAcesso, v. 2, n. 1, p. 107-136, 2008.

MARKETS AND MARKETS. **Conversational AI Market by Component (Platform and Services), Type (IVA and Chatbots), Technology, Application (Customer Support, Personal Assistant, and Customer Engagement and Retention), Deployment Mode, Vertical, and Region** - Global Forecast to 2024. 2019. Disponível em: https://www.marketsandmarkets.com/Market-Reports/conversational-ai-market-49043506.html?gclid=EAIaIQobChMI9Pil4ou15QIVBofVCh3PaQQgEAAYASAAEgKIPvD_BwE. Acesso em: 10 maio 2020.

MARUTITECH. **What is the Working of Image Recognition and How it is Used**? Disponível em: https://marutitech.com/working-image-recognition/. Acesso em: 10 maio 2020.

MASYCHEFF, Alex. **Smart and Smarter: How Structured Content and Ontology Enrich User Experience with Chatbots**. Disponível em: https://intuillion.com/2017/12/12/smart-and-smarter-how-structured-content-and-ontology-enrich-user-experience-with-chatbots/. Acesso em: 20 jul. 2020.

MCCARTHY, John. **What is artificial intelligence?** 2007. Disponível em: http://jmc.stanford.edu/articles/whatisai/whatisai.pdf. Acesso em: 03 maio 2020.

MENDES, Raquel Dias. **Inteligência artificial:** sistemas especialistas no gerenciamento

MENG, Michael; STEINHARDT, Stephanie; SCHUBERT, Andreas. **Application programming interface documentation**: what do software developers want?. Journal of Technical Writing and Communication, v. 48, n. 3, p. 295-330, 2018.

MICROSOFT. **Connect a bot to channels**. 2018. Disponível em: https://docs.microsoft.com/en-us/azure/bot-service/bot-service-manage-channels?view=azure-bot-service-4.0. Acesso em: 22 ago. 2019.

MINDTOUCH. **Chatbot 101**. 2019. Disponível em: https://mindtouch.com/whitepapers/chatbots-101#sizetracker. Acesso em: 22 ago. 2019.

MOURA, T.J.M. **Um chatbot para aquisição automática de perfil do usuário**. 2003. Dissertação (Mestrado em Ciência da Computação) — Universidade Federal de Pernambuco, Recife, Pernambuco, 2003.

MUELLER, Suzana Pinheiro Machado. **Métodos para a pesquisa em Ciência da Informação**. In: Métodos para a pesquisa em Ciência da Informação. 2007.

MUGNAINI, Rogério; STREHL, Letícia. Recuperação e impacto da produção científica na era Google: uma análise comparativa entre o Google Acadêmico e a Web of Science. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, n. Esp, p. 92-105, 2008.

MULDOWNEY, Oisin. **Chatbots**: an introduction and easy guide to making your own. Curses & Magic, 2017.

NOGUEIRA, Fernanda Parolo de Mattos; GRACIOSO, Luciana de Souza; TRIQUES, Maria Lígia; VIEIRA, Claudinei. **Estudo De Usuários Em Biblioteconomia E Ciência Da Informação:** Sistematização Da Produção Científica De Brenda Dervin. In: Ii Encontro Regional Dos Estudantes De Biblioteconomia, Documentação, Ciência E Gestão Da Informação (II EREBD SE/CO/SUL), 2., 2015, São Carlos. Encontro. São Carlos: Ufscar, 2015. p. 0 - 0. Disponível em: http://www.2erebd.ufscar.br/index.php/erebd/erebd/paper/viewFile/62/12. Acesso em: 21 mar. 2020.

OI FUTURO. **Pesquisa de Tendências**: Narrativas para o futuro dos museus: Pesquisa Museus Brasileiros. 2019. Disponível em: http:/oifuturo.org.br/pesquisa-museus-2019. Acesso em: 30 maio 2020.

ORACLE. **Can Virtual Experiences Replace Reality?** 2016. Disponível em: https://www.oracle.com/webfolder/s/delivery_production/docs/FY16h1/doc35/CXResearchVirtualExperiences.pdf. Acesso em: 10 maio 2020.

PAIVA, Fernando. **A história do nascimento da BIA**. 2018. Disponível em: https://www.mobiletime.com.br/noticias/04/05/2018/a-historia-do-nascimento-da-bia/. Acesso em: 31 maio 2020.

PEZZINI, Anderson. **Mineração de textos**: Conceito, processo e aplicações. REAVI-Revista Eletrônica do Alto Vale do Itajaí, v. 5, n. 8, p. 58-61, 2016.

PORTAL ADNEWS (Brasil). **A voz da Arte IBM**: Sistema da IBM faz obra de arte conversar com público. 2017. Disponível em: <http://adnews.com.br/tecnologia/sistema-da-ibm-faz-obra-de-arte-conversar-com-publico.html>. Acesso em: 10 out. 2017.

PORTAL IDGNOW (Org.). **Chatbot brasileiro ajuda a solucionar dúvidas de motoristas sobre carros**. 2017. Disponível em: <http://idgnow.com.br/internet/2016/12/20/chatbot-brasileiro-ajuda-a-solucionar-duvidas-de-motoristas-sobre-carros/>. Acesso em: 10 out. 2017.

RABELLO, Rodrigo. **História dos conceitos e ciência da informação**: apontamentos teórico-metodológicos para uma perspectiva epistemológica. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, v. 13, n. 26, p. 17-46, 2008

RAMALHO, Rogério Aparecido Sá. **Web Semântica**: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação. 2006.

RATTON, Renata. **Piando de orgulho**: projeto da corujinha bia, a nova assistente virtual das bibliotecas puc-rio, conquista prêmio de inovação em bibliotecas elsevier. Projeto da corujinha BIA, a nova assistente virtual das Bibliotecas PUC-Rio, conquista Prêmio de

Inovação em Bibliotecas Elsevier. 2018. Disponível em: http://vrac.puc-rio.br/cgi/cgilua.exe/sys/start.htm?infoid=887&sid=21. Acesso em: 31 maio 2020.

RIBEIRO, Ricardo. **Aplicativo brasileiro traduz manual de carro para o motorista**: sistema permite que Volkswagen virtus 'converse' com seu condutor. Sistema permite que Volkswagen Virtus 'converse' com seu condutor. 2018. Disponível em: https://www1.folha.uol.com.br/mercado/2018/05/aplicativo-brasileiro-traduz-manual-de-carro-para-o-motorista.shtml. Acesso em: 31 maio 2020.

RIBEIRO, Sueliton. **VUI BR**: desenho conversacional. São Paulo: Independente, 2020.

ROCHA, Marco Antônio. **Manual de veículos fica maior, mais indecifrável e dá lugar a aplicativos**. 2018. Disponível em: https://www1.folha.uol.com.br/sobretudo/rodas/2018/04/1964855-pouco-utilizado-manual-de-veiculos-fica-maior-e-ainda-mais-indecifravel.shtml. Acesso em: 30 maio 2020.

RODRIGUES, Bruno César; CRIPPA, Giulia. **A recuperação da informação e o conceito de informação**: o que é relevante em mediação cultural?. Perspectivas em Ciência da Informação, v. 16, n. 1, p. 45-64, 2011.

ROTHERMEL, Alessandra; DE SOUZA DOMINGUES, Maria José Carvalho. MARIA: **Um chatterbot desenvolvido para os estudantes da disciplina "Métodos e Técnicas de Pesquisa em Administração**". SEGET-SIMPÓSIO DE EXCELÊNCIA EM GESTÃO E TECNOLOGIA, Resende, 2007.

RUSSELL, S.; NORVING, P. **Inteligência Artificial**. Elsevier, 2. Ed.: Rio de Janeiro, 2004.

RUSSELL, Stuart J.; NORVIG, Peter. **Artificial intelligence**: a modern approach. Malaysia; Pearson Education Limited,, 2016.

SALESFORCE. **Entenda os Principais Conceitos e o que é Inteligência Artificial**. Disponível em: <https://www.salesforce.com/br/products/einstein/ai-deep-dive/>. Acesso em: 25 jun. 2018.

SARACEVIC, Tefko. **Ciência da Informação: origem, evolução e relações**. Perspectivas em Ciência da Informação, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jul. 1996. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22. Acesso em: 10 mar. 2019.

SCHELLER, Fernando. **Como vai funcionar a Gal, robô de aeroporto da Gol**. 2019. Disponível em: https://economia.estadao.com.br/noticias/geral,como-vai-funcionar-a-gal-robo-de-aeroporto-da-gol,70003055548. Acesso em: 30 maio 2020.

SCHIESSL, M; SHINTAKU, M. **Sistemas de Organização do Conhecimento**. In: Alvares (ORG.). Lillian. Organização da Informação e do Conhecimento: conceitos, subsídios, interdisciplinaridades e aplicações. São Paulo: B4editores, 2012. 248p. Capítulo 2, p. 49/118

SEGURA, M. (2018). **Exemplos de Inteligência Artificial aplicada aos negócios no Brasil**. Disponível em: de https://cio.com.br/exemplos-de-inteligencia- artificial-aplicada-aos-negocios-no-brasil/. Acesso em: 10 mar. 2020.

SETZER, V. W. (1999). **Dado, informação, conhecimento e competência**. DataGramaZero Revista de Ciência da Informação, n. 0, 28.

SEVERINO, Antônio Joaquim. **Metodologia do trabalho científico**. Cortez editora, 2017.

SILVA, Armando Malheiro da. **A informação**: da compreensão do fenómeno e construção do objecto científico. Porto: Afrontamento, 2006.

SILVA, Malcom FB et al. **Desenvolvimento de um Chatbot baseado em Ontologia para Atendimento a Chamados de Suporte ao Cliente**. **Revista Eletrônica de Iniciação Científica em Computação**, v. 17, n. 3, 2019.

SMIT, Johanna W. **A informação na Ciência da Informação**. InCID: Revista de Ciência da Informação e Documentação, Ribeirão Preto, SP, v. 3, n. 2, p. 84-101, jul./dez. 2012. Disponível em: http://www.revistas.usp.br/incid/article/view/48655/52726. Acesso em: 10 mar. 2019.

SOUZA, Renato Rocha et al. **Sistemas de recuperação de informações e mecanismos de busca na web**: panorama atual e tendências. Perspectivas em ciência da informação, v. 11, n. 2, p. 161-173, 2006.

STARTSE. **Startup une chatbot e inteligência artificial para melhorar sua saúde**. 2017. Disponível em: <https://conteudo.startse.com.br/startups/isabela/ta-na-hora/>. Acesso em: 10 out. 2017.

STF. **Vocabulário Jurídico - Tesauros**. Disponível em: http://www.stf.jus.br/portal/jurisprudencia/pesquisarVocabularioJuridico.asp. Acesso em: 10 jul. 2020.

SUBMARINO VIAGENS. **Atendimento**: a voz da arte. A Voz da Arte. Disponível em: https://www.submarinoviagens.com.br/central-de-atendimento.aspx. Acesso em: 04 ago. 2020.

TAKE. **Personalidade do chatbot**: o que é e como defini-la? 2017. Disponível em: https://take.net/blog/chatbots/personalidade-do-chatbot-2. Acesso em: 10 mar. 2020.

TÁLAMO, M. F. G. M.; SMIT, J. W. **Ciência da Informação**: pensamento informacional e integração disciplinar. Brazilian Journal of Information Science, v. 1, n. 1, p. 33-57, 2007. Disponível em: <http://basessibi.c3sl.ufpr.br/brapci/index.php/article/view/0000008749/37fbf34fe79ad82cf215aac68f1833d1>. Acesso em: 05 set. 2017.

TEBA, Sandra. **Finalmente ela chegou! Bia, assistente virtual das bibliotecas puc-rio!** 2017. Disponível em: http://www.dbd.puc-rio.br/wordpress/?p=8885. Acesso em: 31 maio 2020.

THE VERGE. **Baidu launches medical chatbot to help Chinese doctors diagnose patients**. Disponível em: https://www.theverge.com/2016/10/11/13240434/baidu-medical-chatbot-china-melody. Acesso em: 11 out. 2016.

TJSP. **Criadores da Judi, assistente virtual do TJSP, se reúnem com presidente Pereira Calças**. 2019. Disponível em: https://www.tjsp.jus.br/Noticias/Noticia?codigoNoticia=59152&pagina=1. Acesso em: 30 maio 2020.

TOUTAIN, Lídia Maria Batista Brandão (org.). **Para entender a Ciência da Informação**. Salvador: EDUFBA, 2007. p. 13-34. Disponível em: https://repositorio.ufba.br/ri/bitstream/ufba/145/1/Para%20entender%20a%20ciencia%20da%20informacao.pdf. Acesso em: 10 mar. 2019.

TURING, Alan M. **Computing machinery and intelligence**. In: Parsing the turing test. Springer, Dordrecht, 2009. p. 23-65.

UNIPER RESEACH. **Chatbots, a game changer for banking & healthcare, saving $8 billion annually by 2022**. 2017. Disponível em: https://www.juniperresearch.com/press/press-releases/chatbots-a-game-changer-for-banking-healthcare. Acesso em: 10 maio 2020.

WERSIG, Gernot. **Information Science**: the study of postmodern knowledge usage. Information Processing & Management, v. 29, n. 2, p.229-239, 1993. Disponível em: https://ac.els-cdn.com/030645739390006Y/1-s2.0-030645739390006Y-main.pdf?_tid=28fcbc7d-3d48-4edd-8270-ac4be899381e&acdnat=1552491152_0f18d085f71c88dc31fd1e58371e51bd. Acesso em: 21 apr. 2019.

WERTHEIN, Jorge. **A sociedade da informação e seus desafios**. Ciência da informação, v. 29, n. 2, p. 71-77, 2000.

WRIGHT, Alex. **Cataloging the World** - Paul Otlet and the Birth of the information Age. New York: Oxford University Press, 2014.

YAM, Chew-yean. **Emotion Detection and Recognition from Text Using Deep Learning**. Disponível em: https://devblogs.microsoft.com/cse/2015/11/29/emotion-detection-and-recognition-from-text-using-deep-learning/. Acesso em: 25 set. 2015.

ZICK, Laura. **Artificial Intelligence and Libraries:** a primer for librarians. 2006.