



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE GOIÁS

CHARLES SILVA BATISTA
MATHEUS HENRIQUE DE SOUZA

UMA ANÁLISE DO CHATBOT BIA (BRADESCO INTELIGÊNCIA ARTIFICIAL) A
PARTIR DE UM LEVANTAMENTO BIBLIOGRÁFICO.

INHUMAS

2023

CHARLES SILVA BATISTA
MATHEUS HENRIQUE DE SOUZA

UMA ANÁLISE DO CHATBOT BIA (BRADESCO INTELIGÊNCIA ARTIFICIAL) A
PARTIR DE UM LEVANTAMENTO BIBLIOGRÁFICO.

Monografia apresentada à Coordenação do Curso de
Sistemas de Informação do Câmpus Inhumas do Instituto
Federal de Goiás, como parte dos requisitos para
obtenção do título de Bacharel em Sistemas de
Informação. Área de concentração: Metodologias e
Técnicas de Computação.

Orientador: Prof. Dr. Alan Keller Gomes

INHUMAS

2023

Dados Internacionais de Catalogação na Publicação

Batista, Charles Silva

B333 Uma análise do chatbot BIA (Bradesco Inteligência Artificial) a partir de um levantamento bibliográfico [Manuscrito]. / Charles Silva Batista, Matheus Henrique de Souza – Inhumas: IFG, 2023.

77 f.

Bibliografia.

Orientador: Prof. Dr. Alan Keller Gomes.

Trabalho de conclusão de curso de Bacharelado em Sistemas de Informação – IFG/Câmpus Inhumas, 2023.

1. Agentes de conversação. 2. Chatbot. 3. Modelagem conversacional. 4. Souza, Matheus Henrique de. 5. Gomes, Alan Keller (orientador). I. Título.

CDD 003.5

Ficha catalográfica elaborada pela bibliotecária
Larissa Stefane Rodrigues de Lima - CRB/1-3424
Instituto Federal de Educação, Ciência e Tecnologia de Goiás
Câmpus Inhumas – Biblioteca Atena



INSTITUTO FEDERAL
Goiás

MINISTÉRIO DA EDUCAÇÃO

SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA PRÓ-
REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**

**TERMO DE AUTORIZAÇÃO PARA DISPONIBILIZAÇÃO
NO REPOSITÓRIO DIGITAL DO IFG - ReDi IFG**

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia de Goiás, a disponibilizar gratuitamente o documento no Repositório Digital (ReDi IFG), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IFG.

Identificação da Produção Técnico-Científica

- | | |
|--|---|
| ☐ Tese | ☐ Artigo Científico |
| ☐ Dissertação | ☐ Capítulo de Livro |
| ☐ Monografia – Especialização | ☐ Livro |
| ☒ TCC – Graduação | ☐ Trabalho Apresentado em Evento |
| ☐ Produto Técnico e Educacional - Tipo: _____ | |

Nome Completo do Autores: Charles Silva Batista e Matheus Henrique de Souza

Matrícula:

Título do Trabalho: Uma análise do chatbots BIA (Bradesco Inteligência Artificial) a partir de um levantamento bibliográfico

Autorização - Marque uma das opções

1. ☒ Autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso aberto);
2. ☐ Autorizo disponibilizar meu trabalho no Repositório Digital do IFG somente após a data ____/____/____ (Embargo);
3. ☐ Não autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso restrito).

Ao indicar a opção **2 ou 3**, marque a justificativa:

- ☐ O documento está sujeito a registro de patente.
- ☐ O documento pode vir a ser publicado como livro, capítulo de livro ou artigo.
- ☐ Outra justificativa: _____

DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O/A referido/a autor/a declara que:

- i. o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;

- ii. obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia de Goiás os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- iii. cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia de Goiás.

Inhumas, 13/02/2023.

Charles Silva Batista

Matheus Henrique de Souza



INSTITUTO FEDERAL
Goiás

MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE GOIÁS
CÂMPUS INHUMAS

ATA DA SESSÃO PÚBLICA DE APRESENTAÇÃO E DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Aos 9 (nove) dias do mês de fevereiro do ano de 2023, às 14 horas, compareceram perante a banca examinadora os alunos **Charles Silva Batista** e **Matheus Henrique de Souza** autores do Trabalho de Conclusão de Curso - TCC intitulado **Uma análise do chatbot BIA (Bradesco Inteligência Artificial) a partir de um levantamento bibliográfico**, sendo a banca examinadora constituída pelos professores Alan Keller Gomes (orientador), Nisval Ferreira Guimarães (1º examinador) e Rodrigo Cândido Borges (2º examinador) que atribuíram, respectivamente, as seguintes notas: (8,0), (8,0), (8,0), cuja média aritmética é 8,0 (8 pontos), tendo a referida banca considerado os alunos aprovados e aptos para a Colação de Grau de **Bacharel em Sistemas de Informação**.

E por estar conforme, eu Alan Keller Gomes, lavrei a presente ata que vai assinada por mim, pelos membros da banca e pelo coordenador do curso de Bacharelado em Sistemas de Informação.

Orientador(a): Alan Keller Gomes

1º examinador(a): Nisval Ferreira Guimarães

2º examinador(a): Rodrigo Cândido Borges

Coordenador do curso de Bacharelado em Sistemas de Informação: Rodrigo Cândido Borges

Documento assinado eletronicamente por:

- Nisval Ferreira Guimaraes, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 13/02/2023 10:56:00.
- Rodrigo Candido Borges, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 12/02/2023 17:53:49.
- Alan Keller Gomes, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 09/02/2023 16:50:49.

Este documento foi emitido pelo SUAP em 09/02/2023. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifg.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código Verificador: 372380

Código de Autenticação: 515a497b68



Instituto Federal de Educação, Ciência e Tecnologia de Goiás
Av. Universitária, S/Nº, Vale das Goiabeiras, INHUMAS / GO, CEP 75402-556
(62) 3514-9557 (ramal: 9557)

RESUMO

No presente trabalho é apresentada uma análise da ferramenta de diálogo automatizado (*chatbot*) BIA (Bradesco Inteligência Artificial) a partir de um levantamento bibliográfico no qual se destacam os mais recentes avanços sobre a implementação e uso dos *chatbots*. Esses sistemas computacionais inteligentes de conversação são projetados para imitar a conversação humana para permitir orientação e suporte online automatizado aos usuários da ferramenta. Com o aumento dos benefícios gerados para o banco e seus clientes, a adoção da ferramenta foi ampliada para muitos setores por meio do fornecimento da assistência virtual. *Chatbots* utilizam métodos e algoritmos de dois domínios de Inteligência Artificial: Processamento de Linguagem Natural e Aprendizado de Máquina. No entanto, existem muitos desafios e limitações na sua aplicação. No trabalho aqui apresentado é feito um levantamento bibliográfico sobre os avanços recentes no desenvolvimento de *chatbots*, explicitando o uso da Inteligência Artificial e do Processamento de Linguagem Natural. Partindo desse levantamento, é realizada uma análise da BIA, na qual são destacados detalhes de sua implementação, o impacto de sua utilização bem como uma análise das interações geradas a partir da disponibilidade dos recursos de atendimento da BIA para os clientes do banco. Nas considerações finais, são apresentados também os principais desafios e limitações no desenvolvimento e aplicação de ferramentas automatizadas de diálogo tais como a BIA, além de recomendações para pesquisas futuras.

Palavras-chave: agentes de conversação; *chatbot*; *chatbots* sociais; *chatscript*; modelagem conversacional; sistema conversacional; sistema de diálogo humano-computador; sistemas de conversação.

ABSTRACT

This work presents an analysis of the BIA (Bradesco Artificial Intelligence) automated dialogue tool (chatbot) based on a bibliographic survey in which the most recent advances in the implementation and use of chatbots are highlighted. These intelligent conversational computer systems are designed to mimic human conversation to allow automated online guidance and support to tool users. With the increase in benefits generated for the bank and its customers, the adoption of the tool was expanded to many sectors through the provision of virtual assistance. Chatbots use methods and algorithms from two domains of Artificial Intelligence: Natural Language Processing and Machine Learning. However, there are many challenges and limitations in its application. In the work presented here, a review is made of recent advances in the development of chatbots, explaining the use of Artificial Intelligence and Natural Language Processing. Based on this review, an analysis of the BIA is carried out, in which details of its implementation are highlighted, the impact of its use, as well as an analysis of the interactions generated from the availability of the BIA's service resources for the bank's customers. In the final considerations, the main challenges and limitations in the development and application of automated dialogue tools such as BIA are also presented, in addition to recommendations for future research.

Keywords: chatbots; chatscript; conversation systems; conversational agents; conversational modelling; conversational system; human-computer dialogue systema; social chatbots.

LISTA DE ABREVIATURAS E SIGLAS

AIML	<i>Artificial Intelligence Markup Language</i>
ALICE	<i>Artificial Intelligent Internet Computer Entity</i>
AM	Aprendizado de Máquina
BERT	<i>Bidirectional Encoder Representations from transformers</i>
BIA	Bradesco Inteligência Artificial
BLEU	<i>Bilingual Evaluation Understudy</i>
DL	<i>Deep Learning</i>
FAQ	<i>Frequently Asked Questions</i>
GAN	<i>Generative Adversarial Networks</i>
GPT	<i>Generative Pre-trained Transformer</i>
IA	Inteligência Artificial
LSTM	<i>Long Short-Term Memory</i>
METEOR	<i>Metric for Evaluation of Translation with Explicit Ordering</i>
NLU	<i>Natural-Language Understanding</i>
PLN	Processamento de Linguagem Natural
PP	Perplexidade
QE	Quociente Emocional
QI	Quociente Inteligente
RNN	Redes Neurais Recorrentes
SAA	<i>Sensibleness and Specificity Average</i>
XML	<i>Extensible Markup Language</i>

LISTA DE TABELAS

Tabela 1- Resumo do conjunto de dados do chatbot	29
Tabela 2 - Análise de dados das peças	55
Tabela 3 – Cruzamento entre as intenções mapeadas com as hipóteses	59

LISTA DE FIGURAS E IMAGENS

Figura 1 - Tela inicial do aplicativo do Bradesco.....	47
Figura 2 - Tela inicial do menu da BIA.....	48
Figura 3 - Comando "consultar saldo" sem acesso à conta	49
Figura 4 - Comando "fazer um pix" e resposta padronizada	50
Figura 5 - Comando "consultar saldo" com conta acessada	51
Figura 6 - BIA recebendo o comando e processando a solicitação	52
Figura 7 - Tela "Consulta de Saldo" redirecionada pela BIA no App do Banco Bradesco.....	53
Figura 8 - Logotipos da campanha da BIA e logotipo oficial do Bradesco	58

SUMÁRIO

1.	Introdução	10
1.1.	Contextualização	10
1.2.	Motivação e Justificativa	11
1.3.	Objetivos	12
1.4.	Organização do Trabalho	12
2.	Trabalhos de levantamento bibliográfico	13
3.	Materias e métodos	14
4.	Principais aspectos tecnológicos nos chatbots.....	15
4.1.	Histórico dos chatbots	17
4.2.	Abordagens de implementação para chatbots	19
4.2.1.	Chatbots Baseados em Regras	19
4.2.2.	Chatbots de Inteligência Artificial	20
4.3.	Conjunto de dados de Treinamento	27
4.4.	Avaliação	29
4.5.	Aplicação de chatbos	35
5.	Desafios e Limitações no Desenvolvimento dos chatbots.....	36
6.	Estudo de caso: BIA – Bradesco Inteligência Artificial	41
6.1.	História do surgimento da BIA	42
6.2.	Implementação da BIA: Pontos de Destaque.....	44
6.3.	Simulações de uso no aplicativo	47
6.4.	Análise das intenções do Bradesco com uso da BIA	53
6.5.	O Impacto da BIA no segmento bancário	60
7.	Conclusão	62
7.1.	Síntese dos Resultados	62
7.2.	Resultados e Discussões	63
7.2.1.	Implementação de chatbots.....	63
7.2.2.	Aspectos tecnológicos nas implementações	64
7.2.3.	Avaliação das ferramentas	64
7.2.4.	Áreas de aplicação e interesses	65
7.2.5.	Estudo de caso: BIA	66
7.3.	Limitações e Trabalhos Futuros.....	67
	Referências bibliográficas	70

1. INTRODUÇÃO

No presente capítulo é apresentada a introdução do trabalho. Na primeira Seção é feita a contextualização do tema. Em seguida, é apresentada a motivação e a justificativa para realização do estudo aqui relatado. Por fim, os objetivos do estudo e organização do trabalho são apresentados.

1.1. Contextualização

Chatbots são programas inteligentes conversacionais de computador que imitam a conversa humana em sua forma natural (BALA *et al.*, 2017, JIA, 2013; SOJASINGARAYAR, 2020). Esses programas também podem ser definidos como sistema(s) online de diálogo humano-computador com linguagem natural (CAHN, 2017).

Normalmente, são capazes de processar a entrada produzida por um usuário e gerar uma saída processando essa entrada. Eles usam texto em linguagem natural como entrada e a saída deve ser a saída mais relevante para a frase de entrada do usuário (AYANOUZ; ABDELHAKIM; BENHMED, 2020; ALI; KUMAR, 2020).

Portanto, *chatbots* constituem sistemas inteligentes de diálogo automatizado, que podem atender milhares de potenciais usuários de uma só vez. Atualmente, eles são aplicados a uma variedade de campos da vida humana e em aplicações diferentes, desde a área de educação até o comércio eletrônico, abrangendo também saúde, entretenimento etc. Essas ferramentas podem fornecer suporte a seres humanos em diferentes áreas simulando um diálogo entre humanos (CALDARINI; JAF; MCGARRY, 2022; OKUDA; SHODA, 2018).

Um exemplo que está em plena utilização no sistema bancário brasileiro é a BIA (BRADESCO, 2022) (acrônimo de Bradesco Inteligência Artificial), que foi desenvolvida para oferecer atendimento imediato as respostas dos clientes do banco. Criada a partir da plataforma de computação cognitiva Watson, em parceria com a IBM (HIGH, 2012), a Inteligência Artificial (IA) utilizada em sua implementação funciona com base no comportamento do usuário.

Por meio da utilização da BIA é possível que o cliente tire dúvidas, tais como, saber a cotação do dólar, realizar serviços transacionais, por exemplo, pagamento de contas via boleto, transferências bancárias, verificar extratos, faturas pendentes, entre outras possibilidades. Com

a disponibilização e utilização deste sistema conversacional, o banco Bradesco pretende conseguir solucionar dúvidas e problemas comuns dos clientes da forma mais natural possível, facilitando a experiência de uso da ferramenta, 24 horas por dia, 7 dias por semana (ROCKCONTENT, 2018).

1.2. Motivação e Justificativa

Em linhas gerais, *chatbots* são mais atraentes para o usuário do que, por exemplo, uma página estática de Perguntas Frequentes (FAQ) disponível em um endereço eletrônico qualquer. Esses sistemas podem atender simultaneamente vários usuários, sendo, sob essa perspectiva, mais produtivos e menos caros em comparação quando comparados os serviços de suporte ao usuário (cliente) humano (CALDARINI; JAF; MCGARRY, 2022).

Nos últimos anos, com o aumento do poder computacional e o compartilhamento de tecnologias e *frameworks* de código aberto, os sistemas inteligentes de diálogo automatizado se tornaram cada vez mais comuns. Os desenvolvimentos recentes nas técnicas de IA e de Processamento de Linguagem Natural (PLN), tornaram os *chatbots* mais fáceis de implementar, mais flexíveis em termos de aplicação e manutenção e cada vez mais capazes de imitar a conversação humana, apesar de ainda haver imperfeições na interação humano-*chatbot* que precisam ser corrigidas (SHAWAR; ATWELL, 2005) (RAHMAN; AL MAMUN; ISLAM, 2017).

Os *chatbots* são, de fato, menos capazes de entender o contexto conversacional (CHRISTENSEN *et al.*, 2018) e as pistas linguísticas emocionais em comparação com os humanos, o que afeta sua capacidade de conversar de uma maneira mais divertida e amigável (FERNANDES, 2017). Ao mesmo tempo, tendem a assumir papéis tradicionalmente femininos que desempenham com características tradicionalmente femininas e muitas vezes exibindo comportamentos estereotipados, revelando um viés de gênero na implementação e aplicação dessas ferramentas de diálogo automatizado (COSTA, 2018). Por exemplo, o *chatbot* BIA (BRADESCO, 2022) tem estereótipo feminino.

O banco Bradesco disponibilizou, em 2017, uma ferramenta de diálogo automatizado inteligente denominada BIA (BRADESCO, 2022), a qual foi desenvolvida em parceria com IBM e utilização da tecnologia Watson. O intuito do banco é principalmente otimizar processos

já existentes, o uso de dados dentro da organização e inovar na relação entre os clientes e os serviços oferecidos pelo Bradesco. A BIA (BRADESCO, 2022) é um *chatbot* que funciona como uma assistente pessoal para ajudar os clientes nas transações bancárias, além de fornecer dicas de como utilizar o aplicativo, esclarecer dúvidas e operar alguns comandos mais simples. Em 2019 foi atingida a marca de R\$ 100 milhões de interações (SCHORN, 2019).

A motivação para a realização do estudo aqui apresentado vem da percepção da necessidade de compreensão dos principais avanços tecnológicos no desenvolvimento de *chatbots* que estão incorporados em ferramentas como a BIA (BRADESCO, 2022), além de trazer à tona as principais limitações que estão relacionadas ao uso dessas ferramentas.

Através do levantamento bibliográfico, espera-se ser possível identificar os modelos de implementação, métricas e formas de avaliação e também sugerir trabalhos futuros a partir das limitações e problemas encontrados.

1.3. Objetivos

O objetivo principal do presente trabalho é efetuar uma análise da ferramenta de diálogo automatizado (*chatbot*) BIA a partir de um levantamento bibliográfico.

Para atingir o objetivo geral, destacam-se os seguintes objetivos específicos:

1. Efetuar um levantamento bibliográfico sobre as tecnologias envolvidas na implementação de ferramentas de diálogo inteligente;
2. Identificar os aspectos tecnológicos relevantes como abordagens de implementação, conjuntos de dados utilizados na implementação, os principais métodos de avaliação e a aplicação de softwares de diálogo automatizado (*chatbot*) em diferentes domínios;
3. Identificar dos desafios, limitações nos cenários de aplicação de *chatbots*;
4. Analisar a ferramenta BIA a partir do levantamento bibliográfico.

1.4. Organização do Trabalho

O presente trabalho está organizado da seguinte forma: no Capítulo 2 são apresentados detalhes sobre os trabalhos de levantamento bibliográfico; o Capítulo 3 são apresentados os materiais e métodos empregados para a realização do trabalho; o Capítulo 4 apresenta os

principais aspectos tecnológicos; no Capítulo 5 mostra os desafios e limitações ligados ao desenvolvimento dos softwares de diálogo inteligente; no Capítulo 6 é feito um estudo de caso da inteligência artificial do Banco Bradesco (BIA); e no Capítulo 7 a conclusão do trabalho é apresentada.

2. TRABALHOS DE LEVANTAMENTO BIBLIOGRÁFICO

No presente capítulo são apresentados detalhamentos sobre os trabalhos de levantamento bibliográfico, nos quais se baseia a presente pesquisa. Os trabalhos mencionados a seguir tem foco no design e implementação, histórico, métodos de avaliação e a aplicação de *chatbots* em domínio específico.

Abdkul-Kader e Woods (2015) compara técnicas de projeto extraídas de nove artigos selecionados. Os autores se concentram especialmente em *chatbots* vencedores do prêmio *Loebner*¹ e compara modelos usados para desenvolver esses *chatbots* com os modelos apresentados nos artigos selecionados. Os autores discutem áreas em que eles falham e exploram áreas de pesquisa que precisam de atenção.

A pesquisa realizada por (ABDUL-KADER; WOODS, 2015) focou nos *chatbots* baseados em tecnologia na nuvem, programação de *chatbots* e problemas de programação atuais e futuros. Os autores concluem que estabilidade, escalabilidade e flexibilidade são as questões mais importantes para consideração no desenvolvimento.

Cahn (2017) realiza um estudo da literatura sobre o design, arquitetura e algoritmos usados em *chatbots*. Os autores realizaram um estudo de revisão de literatura sistemática e estudo quantitativo. Eles concluíram expressando preocupações acerca da quantidade de material publicado e enfatizaram a importância da interdisciplinaridade.

O estudo de (SINGH *et al.*, 2018) teve como objetivo resolver a questão crítica de identificar técnicas adequadas de Deep Learning (DL). Os autores ofereceram uma visão geral de vários modelos de sistemas de DL para aprendizado. Além disso, forneceram uma visão geral de todo o processo de design, dicas para implementação e links para vários tutoriais, resumos de análise e *pipelines* de DL de código aberto desenvolvidos pela comunidade e modelos pré-treinados. Eles esperavam que esta pesquisa ajudasse na aceleração da adoção do DL em vários domínios científicos. No trabalho de (NURUZZAMAN; HUSSAIN, 2018) é

¹ O Prêmio Loebner foi uma competição anual em inteligência artificial que premiou os programas de computador considerados pelos juízes como os mais humanos.

realizada uma comparação entre as funcionalidades e os requisitos técnicos de onze sistemas de aplicativos de assistente virtual mais comuns.

O estudo realizado por (MOUSSIADES; ADAMOPOULOU, 2020) envolveu duas análises da literatura que discutem a história, tecnologia e aplicações dos *chatbots*. Ao traçar a progressão histórica da ideia inicial até os dias atuais, os autores destacaram potenciais deficiências em cada ponto. Após a apresentação de um esquema de categorização abrangente, os autores discutiram tecnologias de implementação críticas. Por fim, discutiram a arquitetura geral dos assistentes virtuais modernos e as principais plataformas para sua criação. Os autores concluíram que são necessárias mais pesquisas sobre as plataformas de *chatbots* existentes e questões éticas relacionadas a eles.

No trabalho de (CALDARINI; JAF; MCGARRY, 2022) é apresentado (i) um extenso levantamento da literatura bem como o estado da arte em métodos de implementação, com foco em algoritmos de DL, (ii) a identificação dos desafios e limitações nos cenários de aplicação das ferramentas de diálogo automatizado. Eles enfatizam o destaque de que na literatura há uma divisão clara em termos de *chatbots* e sua aplicação. Os algoritmos de DL treinados em grandes conjuntos de dados de domínio, geralmente são implementados como *chatbots* sociais ou complementares. Eles são orientados a tarefas e geralmente são treinados geralmente em domínios menores, específicos e frequentemente confidenciais de conjuntos de dados, assim como são também baseados em abordagens de recuperação de informações ou baseadas em regras, ou uma combinação de ambos.

3. MATERIAS E MÉTODOS

No presente trabalho, como base para o levantamento bibliográfico, foram utilizados o Google Acadêmico e o repositório do *arXiv* para busca dos trabalhos relacionados. Alguns dos termos usados foram: “agentes de conversação”, “chatbots”, “chatbots sociais”, “modelagem conversacional”, “sistema de diálogo humano-computador”, “sistemas de conversação”, e alguns termos na língua inglesa, como “*chat script*”, “*conversation systems*”, “*conversational agents*” e “*social chatbots*”.

Como fonte de Literatura Cinzenta, foi realizada uma busca na plataforma do Google, utilizando termos como: “como funcionam os chatbots”, “BIA do banco Bradesco”, “como lidar com assédio chatbots”, “falsos positivos de chatbots”, dentre outros.

No final, cerca de 107 fontes de pesquisa foram consultadas, entre artigos científicos, monografias e publicações em sites da internet. O foco dos trabalhos levantados se deu principalmente para os materiais publicados de 2007 a 2021. Os títulos dos trabalhos consultados no Google Acadêmico foram filtrados em sua maioria de acordo com a relevância para com os objetivos deste trabalho.

Os materiais e métodos empregados na pesquisa aqui apresentada são:

- 1) No levantamento bibliográfico sobre as tecnologias envolvidas na implementação de *chatbots* foram consultados os trabalhos de revisão sistemática de literatura (ABDUL-KADER; WOODS, 2015; CAHN, 2017; CALDARINI; JAF; MCGARRY, 2022; MOUSSIADES; ADAMOPOULOU, 2020; NURUZZAMAN; HUSSAIN, 2018; SINGH *et al.*, 2018).
- 2) Na identificação dos principais aspectos tecnológicos das ferramentas de diálogo automatizado são replicados os elementos apresentados em (CALDARINI; JAF; MCGARRY, 2022) por ser o trabalho mais recente disponível no momento;
- 3) Na identificação dos desafios, limitações nos cenários de aplicação de *chatbots* optou-se pelo detalhamento apresentado em (CALDARINI; JAF; MCGARRY, 2022) por ser o mais recente no momento;
- 4) Na análise da ferramenta de diálogo automatizado BIA foi feito um levantamento na literatura cinzenta, conforme (GOMES; RÉBEL; ROCHA, 2000), sendo consultados materiais disponíveis em blogs, sites, manuais e relatórios.

4. PRINCIPAIS ASPECTOS TECNOLÓGICOS NOS CHATBOTS

No presente capítulo são apresentados os principais aspectos tecnológicos relacionados com ferramentas de diálogo automatizado (ou inteligente) que são:

- **Histórico e Evolução dos *Chatbots*:** este aspecto engloba todos os trabalhos que apresentam uma descrição detalhada da evolução ao longo do tempo. Essa categoria é

fundamental pois ajuda a entender as tendências e tecnologias que ascenderam ou foram descartadas ao longo do tempo, indicando sua evolução. Também ajudou a descobrir como e por que esses assistentes virtuais e como suas aplicações e propósitos mudaram ao longo do tempo. O subcapítulo 4.1 oferece uma visão geral sobre a história e evolução.

- **Implementação de *Chatbots*:** este aspecto inclui artigos que apresentam exemplos de projeto arquitetônico e implementação. Esta categoria permitiu identificar os algoritmos mais comuns e os algoritmos específicos que são usados para diversos tipos de *chatbots* com base na finalidade da aplicação. Isso também permitiu identificar os padrões da indústria em termos de modelos e algoritmos, bem como suas deficiências e limitações. Abordagens de implementação detalhadas são fornecidas no subcapítulo 4.2.
- **Conjunto de dados:** esta categoria foi usada para classificar *chatbots* de acordo com o conjunto de dados usado para treinar algoritmos de aprendizado de máquina para o desenvolvimento de modelo de linguagem. O subcapítulo 4.3 destaca os principais conjuntos de dados que foram usados em estudos anteriores.
- **Avaliação de *Chatbots*:** Para este aspecto, alguns artigos focaram nos métodos de avaliação e métricas utilizadas para medir o desempenho dos *chatbots*. Foi importante identificar esses artigos para entender a forma que eles são avaliados e as métricas e métodos de avaliação utilizados. As várias métricas de avaliação são apresentadas no subcapítulo 4.4.
- **Aplicações de *Chatbots*:** este aspecto engloba todos os exemplos de *chatbots* aplicados a um domínio específico, como educação, finanças, suporte ao cliente e psicologia. Os artigos pertencentes a esta categoria ajudaram a vincular informações de categorias anteriores e a entender melhor quais modelos e quais recursos são usados para quais aplicativos para atender a diferentes propósitos. No subcapítulo 4.5 é apresentada uma visão geral sobre suas aplicações.

Essas categorias surgiram do levantamento bibliográfico. Os artigos revisados abrangeram uma ou mais dessas categorias, muitas vezes indicadas no título ou no resumo. O desenvolvimento de aplicações de *chatbot* requer, de fato, um estudo de diferentes implementações e outras aplicações (que são o resultado da evolução dos *chatbots* ao longo do tempo), bem como um conjunto de dados e um modelo de avaliação.

4.1. Histórico dos chatbots

Embora a busca por ser capaz de criar algo que possa entender e se comunicar com seu criador tenha raízes profundas na história humana, Alan Turing é considerado a primeira pessoa a conceituar a ideia de um *chatbot* em 1950, quando propôs a questão: “As máquinas podem pensar?”. A descrição de Turing do comportamento de uma máquina inteligente evoca o conceito comumente entendido de *chatbot* (TURING, 2009).

Os assistentes virtuais evoluíram com o aumento progressivo das capacidades computacionais e avanços nas ferramentas e técnicas de Processamento de Linguagem Natural. A primeira implementação de um *chatbot*, que dependia fortemente de regras linguísticas e técnicas de correspondência de padrões, foi concebido em 1966 com o desenvolvimento do ELIZA (WEIZENBAUM, 1996).

Esse *chatbots* foi o primeiro programa para processamento de linguagem natural da história e foi criado por Joseph Weizenbaum no laboratório de Inteligência Artificial do Massachusetts Institute of Technology (MIT). Ele pôde se comunicar com o usuário através do programa de correspondência de palavras-chave. Ele busca uma regra de transformação apropriada para reformular a entrada e fornecer uma saída, ou seja, uma resposta ao usuário. Também foi um sistema de referência que estimulou mais pesquisas na área. No entanto, o escopo de conhecimento era limitado porque dependia da identificação mínima do contexto e, geralmente, as regras de correspondência de padrões não são flexíveis para serem facilmente implementadas em novos domínios (LI; HE; SHUM, 2018; WEIZENBAUM, 1996; ZEMČÍK, 2019).

Uma evolução marcante nos *chatbot* na década de 1980 se deu devido ao uso de Inteligência Artificial. O sistema ALICE (*Artificial Intelligent Internet Computer Entity*), baseado na *Artificial Intelligence Mark up Language* (AIML), que é uma extensão do XML, foi desenvolvido especialmente para que o conhecimento do padrão de diálogo pudesse ser adicionado ao software do ALICE (ATWELL; ABUSHAWAR, 2015) para expandir sua base de conhecimento. Os objetos de dados no AIML são compostos de tópicos e categorias. As categorias são a unidade básica de conhecimento, que são compostas por uma regra para combinar as entradas do usuário com as saídas do *chatbot*. A entrada do usuário é representada por padrões de regras, enquanto a saída do *chatbot* é definida pelo modelo de regra, base de conhecimento ALICE. A adição de novos objetos de dados em AIML representou uma melhoria significativa em sistemas anteriores de correspondência de padrões, pois a base de

conhecimento era facilmente expansível. Além disso, o *Chatscript*, o sucessor do AIML, também foi a tecnologia base por trás de outros *chatbots* premiados do Prêmio Loebner. A ideia principal por trás dessa tecnologia inovadora era combinar entradas textuais de usuários a um tópico, e cada tópico teria uma regra específica associada a ele para gerar uma saída. O *Chatscript* inaugurou uma nova era para a evolução da tecnologia dos assistentes virtuais. Começou a mudar o foco para a análise e compreensão semântica (ATWELL; ABUSHAWAR, 2015; CAHN, 2017; LI; HE; SHUM, 2018; MLADENIĆ; BRADEŠKO, 2012; WILCOX; WILCOX, 2014; ZEMČÍK, 2019).

A principal limitação em depender de regras e correspondência de padrões em assistentes virtuais é que eles são dependentes de domínio, o que os torna inflexíveis, pois dependem de regras escritas manualmente para domínios específicos. Com os recentes avanços em técnicas de Aprendizado de Máquina (AM) e ferramentas de PLN aliados à disponibilidade de poder computacional, novos frameworks e algoritmos foram criados para implementar *chatbots* “avançados” que não se baseiam em regras e técnicas de correspondência de padrões, o que incentivou o uso comercial deles. A aplicação de algoritmos de aprendizado de máquina em *chatbots* tem sido investigada e novas arquiteturas surgiram (RAHMAN; AL MAMUN; ISLAM, 2017; SHAWAR; ATWELL, 2005).

A aplicação de *chatbots* se expandiu com o surgimento de algoritmos de DL. Uma das novas e mais interessantes aplicações é o desenvolvimento de assistentes pessoais inteligentes (como Alexa da Amazon, Siri da Apple, Google Assistant do Google, Cortana da Microsoft e Watson da IBM). Os assistentes pessoais ou agentes de conversação que normalmente podem se comunicar com o usuário por meio de voz geralmente são integrados em *smartphones*, *smartwatches*, alto-falantes e monitores domésticos dedicados e até carros. Por exemplo, quando o usuário fala uma palavra ou frase de ativação, o dispositivo é ativado e o assistente pessoal inteligente começa a ouvir. Por meio do Natural Language Understanding (NLU) o assistente pode entender comandos e responder às solicitações do usuário, geralmente fornecendo informações (por exemplo, “Alexa, qual é o clima hoje em Los Angeles?” Em Los Angeles o tempo está ensolarado e está 24°C”), ou concluindo tarefas (por exemplo, “Ok Google, toque minha lista de reprodução matinal no Spotify”). No entanto, a tarefa de entender a linguagem humana provou ser bastante desafiadora devido às variações tonais, regionais, locais e até pessoais na fala humana.

Todos os assistentes pessoais inteligentes apresentam as mesmas características essenciais em termos de tecnologias utilizadas, interface de usuário e funcionalidades. Alguns

chatbots têm, no entanto, uma personalidade mais desenvolvida do que outros, e os mais desenvolvidos também podem proporcionar entretenimento e não apenas assistência nas tarefas do dia-a-dia; esses são chamados de *chatbots* sociais. Um exemplo interessante de um assistente virtual social é o XiaoIce da Microsoft. XiaoIce foi feito para ser um companheiro de longo prazo para o usuário e, para alcançar um alto envolvimento com o usuário, foi projetado para ter uma personalidade, um Quociente Inteligente (QI) e um Quociente Emocional (QE). Modelagem de conhecimento e memória, compreensão de imagem e linguagem natural, raciocínio, geração e previsão são exemplos de recursos de QI. Estes são componentes críticos do desenvolvimento de habilidades de diálogo. Eles são necessários para que os *chatbots* sociais atendam às necessidades específicas dos usuários e os ajudem. A habilidade mais crítica e sofisticada é o *Core Chat*, que pode se envolver em conversas longas e de domínio aberto com os usuários. A empatia e as habilidades sociais são dois componentes críticos do QE. O mecanismo de conversação do XiaoIce usa um gerenciador de diálogo para acompanhar o estado da conversa e seleciona o *Core Chat* (o Componente Generativo de domínio aberto) ou a habilidade de diálogo para gerar uma resposta. Portanto, o modelo incorpora recursos de recuperação de informações e generativos (DORMEHL, 2018; SPENCER, 2018; ZHOU *et al.*, 2020).

4.2. Abordagens de implementação para chatbots

Nesta seção, uma visão geral dos métodos de implementação dos *chatbots* é apresentada. É dada também uma distinção entre os que são baseados em regras e outros baseados em Inteligência Artificial. Dentro dos baseados em IA, uma comparação entre os de recuperação de informações e generativos também é apresentada, assim como as desvantagens e limitações de cada abordagem de implementação e suas melhorias recentes.

4.2.1. Chatbots Baseados em Regras

As primeiras tentativas de implementação de *chatbots* foram baseadas em regras. Os modelos baseados em regras geralmente são mais fáceis de projetar e implementar, mas são limitados em termos de capacidades, pois apresentam dificuldades para responder a consultas complexas. Os assistentes virtuais baseados em regras respondem às consultas dos usuários procurando correspondências de padrões; portanto, é provável que produzam respostas

imprecisas quando se deparam com uma frase que não contém nenhum padrão conhecido. Além disso, codificar manualmente as regras de correspondência de padrões pode ser difícil e demorado. Ainda, as regras de correspondência de padrões são frágeis, altamente específicas de domínio e não transferem bem de um problema para outro.

4.2.2. Chatbots de Inteligência Artificial

Os modelos de IA, ao contrário dos modelos baseados em regras, são baseados em algoritmos de aprendizado de máquina que permitem aprender com um banco de dados existente de conversas humanas. Para isso, eles precisam ser treinados por meio de algoritmos de *Machine Learning* que podem treinar o modelo usando um conjunto de dados de treinamento. Por meio do uso de algoritmos de *Machine Learning*, não há mais a necessidade de definir e codificar manualmente novas regras de correspondência de padrões, o que permite que os *chatbots* sejam mais flexíveis e não mais dependentes de conhecimentos específicos do domínio. Como afirmado, os modelos de IA podem ser categorizados em modelos baseados em recuperação de informações e modelos generativos.

4.2.2.1. Modelos de Recuperação de Informação

Modelos baseados em Recuperação de Informações são projetados para que, dado um conjunto de dados de informações textuais, o algoritmo seja capaz de recuperar as informações necessárias com base na entrada do usuário. O algoritmo usado geralmente é um algoritmo de aprendizado superficial; no entanto, também há casos de modelos de Recuperação de Informação que utilizam algoritmos baseados em regras de *Deep Learning*.

Os modelos baseados em Recuperação de Informações incluem um conjunto predefinido de respostas possíveis; o *chatbot* processa a consulta do usuário e com base nessa entrada ele escolhe uma das respostas disponíveis em seu conjunto. A base de conhecimento para esse tipo de modelo geralmente é formada por um banco de dados de pares de perguntas e respostas. Um índice de bate-papo é construído a partir desse banco de dados, a fim de listar todas as respostas possíveis com base na mensagem que as solicitou. Quando o usuário fornece uma entrada ao *chatbot*, o assistente virtual trata essa entrada como uma consulta, e um modelo de recuperação de informações semelhante aos usados para consultas na web é usado para corresponder a entrada do usuário a outras parecidas no índice do *chat*. A saída devolvida ao

usuário é, portanto, a resposta emparelhada com a pergunta selecionada entre as presentes no índice do *chat* (LI; HE; SHUM, 2018). A principal vantagem deste modelo é que garante a qualidade das respostas, uma vez que não são geradas automaticamente. Esses modelos viram um aumento na popularidade com o advento da Web 2.0 e o aumento da informação textual disponível que pode ser recuperada em plataformas de mídia social, fóruns e *chats* (YAN; SONG; WU, 2016).

Uma das principais desvantagens dessa abordagem é que criar a base de conhecimento necessária pode ser caro, demorado e tedioso. Além disso, se o grande volume de dados disponível fornecer um maior conjunto de treinamento e uma base de conhecimento mais ampla, também implicará que será ainda mais desafiador combinar a entrada de um usuário com a resposta correta; uma quantidade significativa de tempo e recursos devem ser empregados para treinar o sistema para selecionar uma das respostas corretas disponíveis (YAN; SONG; WU, 2016).

Finalmente, os sistemas de Recuperação de Informação, devido ao fato de não gerarem respostas, mas sim recuperarem respostas de um conjunto pré-definido em sua base de conhecimento, são indiscutivelmente menos adequados para serem usados como algoritmo subjacente para agentes de conversação ou bate-papo, os chamados *chatbots* sociais. Modelos de Recuperação de Informação são de fato menos adequados para desenvolver uma personalidade, o que é uma característica importante para esse tipo de assistente virtual (LI; HE; SHUM, 2018).

No entanto, progressos foram feitos no desenvolvimento de novos algoritmos de Recuperação de Informação nos últimos tempos, e vale a pena mencionar quais algoritmos de *Machine Learning* estão sendo usados atualmente como tecnologia subjacente para esse tipo de modelo. (LU; LI, 2013) propôs um novo modelo para representar coocorrência textual local e mapear informações hierárquicas entre domínios para termos semanticamente mais distantes. Esse modelo foi baseado na ideia de que quanto maior a coocorrência de dois termos entre domínios, mais intimamente relacionados são os dois termos. Assim, uma alta coocorrência dentro de um domínio específico poderia disponibilizar o processo de recuperação da informação. Esse modelo foi, portanto, baseado em duas etapas: modelagem de tópicos para texto paralelo e obtenção da arquitetura hierárquica. O primeiro passo visa encontrar padrões de coocorrência significativos de palavras.

A segunda etapa visa modelar a arquitetura de coocorrências entre os tópicos. Essa arquitetura será usada para criar a rede neural que alimenta esse algoritmo de aprendizado de

máquina. O desenvolvimento interessante desse modelo reside, portanto, no uso de coocorrências de palavras para definir um contexto. O objetivo subjacente desta pesquisa foi usar informações contextuais para melhorar o desempenho de correspondência para modelos de recuperação de informação (LU; LI, 2013).

Um desenvolvimento interessante, que visa levar em consideração o turno anterior na conversa, obtendo assim mais informações contextuais para melhorar a qualidade e a correção da saída, é o proposto por (YAN; SONG; WU, 2016). Nesse modelo, o processo de recuperação de informações é aprimorado por uma *Deep Neural Network* que classifica não apenas o par de pergunta/resposta correspondente à entrada do último usuário, mas também os pares de pergunta/resposta que correspondem a versões reformuladas de turnos de conversa anteriores. As listas de classificação correspondentes a diferentes reformulações são então mescladas. Dessa forma, informações contextuais podem ser aproveitadas a partir de consultas anteriores do usuário e essas informações podem ser usadas para recuperar uma resposta melhor dentro da base de conhecimento (YAN; SONG; WU, 2016).

4.2.2.2. Modelos Geradores

Modelos de base generativa, como o nome sugere, geram novas respostas palavra por palavra, com base na entrada do usuário. Esses modelos são, portanto, capazes de criar frases inteiramente novas para responder às consultas dos usuários; no entanto, eles precisam ser treinados para aprender a estrutura e a sintaxe das frases, e as saídas podem ter alguma falta de qualidade ou consistência (SHANG; LU; LI, 2015; SORDONI *et al.*, 2015; SUTSKEVER; VINYALS; LE, 2014; VINYALS; LE, 2015).

Os modelos generativos geralmente são treinados em um grande conjunto de dados de frases naturais emitidas a partir de uma conversa. O modelo aprende a estrutura da frase, a sintaxe e o vocabulário por meio dos dados que foram alimentados. O objetivo geral é que o algoritmo seja capaz de gerar uma resposta apropriada e linguisticamente correta com base na sentença de entrada. Esta abordagem é geralmente baseada em um algoritmo de *Deep Learning* composto por um modelo de Rede Neural *Encoder-Decoder* com mecanismos de *Long-Short-Term-Memory* para contrabalançar o efeito gradiente de fuga presente em Redes Neurais Recorrentes tradicionais (VINYALS; LE, 2015).

4.2.2.3. Algoritmos Padrão da Indústria

Entre os modelos de IA, os modelos *Sequence to Sequence* se tornaram o padrão do setor para modelagem de *chatbot*. Eles foram introduzidos pela primeira vez para resolver problemas de tradução automática, mas os princípios subjacentes de fato parecem ter um bom desempenho também para a geração de linguagem natural. Esses modelos são compostos por duas Redes Neurais Recorrentes (RNN), um *Encoder* e um *Decoder*. A frase de entrada do usuário do *chatbot* se torna a entrada do *Encoder*, que processa uma palavra por vez em um estado oculto específico da RNN. O estado final representa a intenção da sequência e é chamado de vetor de contexto. O decodificador toma o vetor de contexto como sua entrada e gera outra sequência (ou sentença) uma palavra por vez.

O objetivo geral para este modelo probabilístico é aprender a gerar a resposta mais provável dado o contexto conversacional, que neste caso é constituído pelo turno anterior na conversa, ou a sentença de entrada. Na fase de aprendizado, a resposta, ou sentença de saída, é dada ao modelo para que ele possa aprender por retro propagação. Para a fase de interferência, duas abordagens diferentes podem ser usadas. A abordagem de busca de feixe fornece vários candidatos como sentença de entrada e a sentença de saída é selecionada com base na probabilidade mais alta. Uma abordagem mais gananciosa usa o *token* de saída previsto como entrada para prever a próxima sentença da conversa (VINYALS; LE, 2015).

Este modelo oferece algumas vantagens interessantes. Primeiro, não envolve conhecimento específico de domínio, mas é uma solução de ponta a ponta que pode ser treinada usando diferentes conjuntos de dados, portanto, em diferentes domínios. Além disso, embora o modelo não precise de conhecimento específico do domínio para fornecer resultados valiosos, ele pode ser adaptado para trabalhar com outros algoritmos se for necessária uma análise mais aprofundada do conhecimento específico do domínio. É, portanto, um modelo simples, mas amplamente geral e flexível que pode ser usado para resolver diferentes tarefas de PLN (LI; HE; SHUM, 2018; VINYALS; LE, 2015). Por essas razões, o modelo *Sequence-to-Sequence* parece ter se tornado a escolha padrão da indústria para geração de diálogo e muitas tarefas de PLN recentemente.

No entanto, tem um limite considerável: a totalidade das informações contidas na sentença de entrada deve ser codificada em um vetor de comprimento fixo, o vetor de contexto, e assim, quanto maior a sentença, mais informações se perdem no processo. É por isso que os modelos *Sequence to Sequence* não funcionam bem quando devem responder a frases mais longas e tendem a dar respostas vagas. Além disso, ao gerar uma resposta, esses modelos

tendem a se concentrar em uma única resposta, o que cria uma falta de coerência nos turnos de uma conversa (JURAFSKY; MARTIN, 2008; SOJASINGARAYAR, 2020; STRIGÉR, 2017).

4.2.2.4. Transformers

É claro que uma das inovações mais interessantes em modelos de linguagem de *Deep Learning* foi a introdução de *Transformers*, apresentada pela primeira vez por (VASWANI *et al.*, 2017), no jornal “*Attention is all you need*” (Atenção é tudo que você precisa). *Transformers* são modelos de linguagem baseados exclusivamente no mecanismo *Attention*. Os transformadores são hoje o modelo de escolha para os desafios de PLN, substituindo modelos RNN como memória de longo prazo (LSTM) por pesar diferencialmente a relevância de cada parte dos dados de entrada. Além disso, eles fornecem paralelização de treinamento que permite o treinamento em conjuntos de dados maiores do que era originalmente alcançável.

Isso levou ao desenvolvimento de sistemas pré-treinados como BERT (*Bidirectional Encoder Representations from transformers*) (DEVLIN *et al.*, 2018) e GPT (*Generative Pre-trained Transformer*), que foram treinados com enormes conjuntos de dados de linguagem, como *Wikipedia Corpus* e *Common Crawl*, e podem ser ajustados para aplicações específicas. Desde então, várias versões diferentes do *Transformer* foram apresentadas, como o *Reformer*) e o *Transformer XL* (DAI *et al.*, 2019). Cada versão do transformador foi desenvolvida para responder a desafios específicos para a tarefa em mãos. Embora os transformadores tenham sido introduzidos para responder a desafios de Tradução por Máquinas, eles podem ser adaptados e modificados para realizar tarefas de modelagem de diálogo.

Em (DAI *et al.*, 2019), os autores propõem uma versão atualizada do *Transformer* denominada *Transformer XL*. Esse modelo pode ir além das limitações de contexto de comprimento fixo do *Transformer*, usando a recorrência em nível de sentença. Os transformadores mostram um potencial de aprendizado de dependência de longo prazo, mas são limitados pelo contexto de comprimento fixo no cenário de modelagem de linguagem. Os autores apresentam uma arquitetura neural única chamada *Transformer-XL* que permite aprender a dependência além de um determinado comprimento sem quebrar a coerência temporal. Ele compreende um mecanismo de recorrência em nível de segmento e uma técnica de codificação posicional exclusiva. Esta solução visa capturar a dependência de longo prazo e resolver o problema de fragmentação de contexto. Embora essa abordagem ainda não tenha sido aplicada à modelagem de diálogo, pode-se argumentar que, uma vez implementada a modificação apropriada e necessária, ela pode ser útil para superar alguns dos problemas que os modelos de diálogo atuais apresentam, isto é, a compreensão do contexto.

4.2.2.5. Reformer

Em (KITAEV; KAISER; LEVSKAYA, 2020), os autores apresentam o *Reformer*, uma versão mais eficiente do *Transformer*, que faz uso de duas técnicas para melhorar o *Transformer* em termos de eficiência. Em primeiro lugar, os autores substituem a atenção do produto escalar por uma que emprega *hashing* sensível à localidade, aumentando sua complexidade de $O(L^2)$ para $O(L \log L)$, onde L é o tamanho da sequência. Em segundo lugar, eles empregam camadas residuais reversíveis em vez dos resíduos padrão, o que permite armazenar a ativação apenas uma vez no processo de treinamento de N vezes, onde N é o número de camadas. Como resultado, o *Reformer* é significativamente mais eficiente em termos de memória e substancialmente mais rápido em sequências mais longas.

4.2.2.6. Meena

Em (ADIWARDANA *et al.*, 2020), os autores apresentam o Meena, um modelo de *chatbot* generativo que foi treinado de ponta a ponta em 40 bilhões de palavras extraídas e filtradas de discussões de mídia social de domínio público. Com Meena, os autores estendem os limites da abordagem de ponta a ponta para mostrar que um modelo de baixa perplexidade em grande escala pode produzir saídas de linguagem de qualidade.

Os autores empregam um modelo seq2seq (BAHDANAU; CHO; BENGIO, 2014; SUTSKEVER; VINYALS; LE, 2014) com o *Evolved Transformer* (SO; LE; LIANG, 2019) como arquitetura principal. As quatro características mais importantes da arquitetura do *Evolved Transformer* são a utilização de (i) grandes convoluções separáveis em profundidade, (ii) Unidades lineares fechadas (DAUPHIN *et al.*, 2017), (iii) estruturas ramificadas e (iv) ativações *swish* (RAMACHANDRAN; ZOPH; LE, 2017). Tanto o codificador quanto o decodificador do *Evolved Transformer* geraram independentemente uma porção inferior ramificada com amplas convoluções. Também em ambas as situações, a porção posterior é essencialmente idêntica ao *Transformer* (SO; LE; LIANG, 2019). O modelo é treinado em diálogos de múltiplas voltas onde a sentença de entrada é composta por todas as voltas do contexto (até 7) e a sentença de saída é a resposta.

Para quantificar a qualidade do Meena e compará-lo com outros *chatbots*, (ADIWARDANA *et al.*, 2020) também se propõe uma métrica de avaliação humana eficaz. *Sensibleness and Specificity Average* (SSA), em português Sensibilidade e Especificidade

Média, incorpora duas qualidades básicas de um *chatbot* humano: fazer sentido e ser específico em suas respostas. Os autores pedem que juízes humanos avaliem as respostas com base nesses dois aspectos.

4.3. Conjunto de dados de Treinamento

Serão apresentados neste capítulo os conjuntos de dados usados mais frequentemente para treinar modelos de *chatbots* com aprendizado profundo. Primeiramente, deve-se diferenciar os conjuntos de dados de domínio público e de domínio fechado. Na literatura, parece haver alguns conjuntos de dados de domínio que são mais comumente usados: OpenSubtitles, Cornell e conjunto de dados do DailyDialog.

Cada um desses conjuntos de dados será apresentado de forma mais detalhada. Por outro lado, parece não haver conjuntos de dados de domínio fechado usados com frequência na literatura pesquisada. Isso pode ser devido à especificidade de tais conjuntos de dados, que geralmente atendem a necessidades e objetivos específicos dentro de um escopo definido. Para não prolongar, não será discutido cada um desses conjuntos de dados em detalhes.

- **OpenSubtitles:** usado por vários trabalhos na literatura. É um conjunto de dados de domínio aberto feito de legendas de filmes no formato XML. O banco de dados oferece acesso aberto a legendas de filmes e séries de vários idiomas. O conjunto de dados parece ser bastante grande, contendo milhões de pares de frases, porém os dados não parecem ser dos melhores qualidade; descrições de cenas, legendas e frases segmentadas aparecem entre os diálogos, o que pode ser problemático ao treinar um *chatbot* de domínio aberto, uma vez que a coesão do diálogo pode ser perdida (CHRISTENSEN *et al.*, 2018; KLEIN *et al.*, 2017; LI *et al.*, 2016, 2017; LISON; TIEDEMANN, 2016; VINYALS; LE, 2015; ZHONG; WANG; MIAO, 2019).
- **Cornell:** também é bastante usado para treinar o sistema de diálogo de domínio aberto. Ele apresenta uma coleção rica em metadados de conversas fictícias tomadas a partir de roteiros de filmes brutos, consistindo em mais de 300.000 enunciados totais. Os dados são apresentados no formato “txt”. Como os dados foram retirados de roteiros de filmes, tendem a ser de boa qualidade. No entanto, o conjunto de dados pode não ser grande o suficiente para treinar modelos de linguagem (DANESCU-NICULESCU-MIZIL; LEE, 2011; HE *et al.*, 2019; ROLLER *et al.*, 2020; ZHONG; WANG; MIAO,

2019). (Nenhuma informação foi fornecida pelos autores sobre o número de tokens no conjunto de dados. Por esse motivo, todo o conjunto de dados foi transformado usando o processador de dados *keras.preprocessing.text*. Os resultados são mostrados na Tabela 1).

- **DailyDialog:** este conjunto de dados é feito por humanos e consiste em várias conversas de múltiplas voltas. São um pouco mais de 13.000 conversas, cada uma composta por 8 turnos por orador na média. Os dados são anotados: as anotações contêm informações sobre emoções e intenções. Essas informações podem ser úteis no treinamento de alguns modelos de linguagem. Embora os dados pareçam mais limpos em comparação com outros conjuntos de dados, é também menor, portanto, menos adequado para modelos de diálogo mais complexos (HE *et al.*, 2019; KIM *et al.*, 2019; LI *et al.*, 2017; ZHONG; WANG; MIAO, 2019). (Nenhuma informação foi fornecida pelos autores sobre o número de tokens no conjunto de dados. Por esse motivo, todo o conjunto de dados foi transformado usando o processador de dados *keras.preprocessing.text*. Os resultados são mostrados na Tabela 1).

Tabela 1- Resumo do conjunto de dados do chatbot

Conjunto de dados	Tipo de conteúdo e fonte	Nº de Frases	Tokens	Fonte
OpenSubtitles	Legenda de filmes. Todo o banco de dados do repositório do OpenSubtitles.org	441.5 M (publicado em 2018)	3.2 G (publicado em 2018)	(LISON, 2016)
Cornell	Scripts brutos de filmes. Conversas fictícias extraídas de scripts de filmes	304,713	48,177	(DANESCU-NICULESCU-MIZIL; LEE, 2011)
DailyDialog	Diálogos para aprendizes de Inglês. Dados brutos rastreados de vários websites que oferecem conteúdo para quem está aprendendo Inglês	103,632 (13,118 diálogos com 7.9 voltas cada em média)	17,812	(LI; <i>et al.</i> , 2017)

Fonte: Caldarini, Jaf e Mcgarry (2022).

4.4. Avaliação

Avaliar sistemas de diálogo tem se mostrado uma tarefa desafiadora, uma vez que a conversação humana atende a diferentes objetivos e funções. Dependendo do objetivo do *chatbot*, as métricas usadas para avaliar o diálogo podem mudar. Um *chatbot* de assistente pessoal será avaliado principalmente com base na eficácia da interação (executou a tarefa que o usuário pediu? A troca foi eficiente?), enquanto um complementar será avaliado em sua capacidade de manter a conversa e envolver os usuários. Existem duas principais formas de avaliá-los: avaliação humana e métricas de avaliação automatizada. A avaliação humana consiste em pedir a um grupo de participantes que interaja com o assistente virtual e, em seguida, avalie os diferentes aspectos da interação de acordo com um quadro de avaliação ou questionários.

Os participantes geralmente avaliam os diferentes aspectos da interação com base em uma escala que pode ser usada para traçar médias e medir o desempenho de qualidade em termos de eficiência, eficácia e satisfação dos usuários (BENTON; RADZIWILL, 2017). Embora a avaliação humana permita avaliar a qualidade de diferentes aspectos da interação, é

dispendiosa (uma vez que há necessidade de alocar recursos humanos para a avaliação), consome muito tempo, não facilmente escalável e sujeito a vieses (mesmo ao seguir uma estrutura de avaliação, diferentes as pessoas podem avaliar a mesma interação de forma diferente).

No entanto, a avaliação humana pode levar em consideração vários aspectos da troca e avaliar a conversa em diferentes níveis; além disso, o quadro de avaliação pode ser adaptado com base no objetivo principal e funções do *chatbot* ou sistema de diálogo. Por esses motivos, as métricas de avaliação humana são usadas em várias partes das literaturas analisadas, como (CHRISTENSEN *et al.*, 2018; SORDONI *et al.*, 2015). O PARAdigm for DIAlogue System Evaluation é um dos quadros de avaliação mais amplamente utilizados para combinar diferentes níveis de avaliação (PARADISE). Em primeiro lugar, PARADISE avalia variáveis subjetivas como (i) usabilidade do sistema, (ii) clareza, (iii) naturalidade, (iv) simpatia, (v) contornar mal-entendidos e (vi) vontade de usar o sistema novamente. Ele alcança isso solicitando *feedback* do usuário por meio de questionários. Em segundo lugar, através da otimização do sucesso das tarefas e diminuindo os custos do diálogo, o PARADISE visa quantificar a eficácia do *bot* (CAHN, 2017; HUNG *et al.*, 2009; KEMPSON; GARGETT; GREGOROMICHELAKI, 2007; LI; WESTON; ROLLER, 2019; LIN *et al.*, 2020; SEDOC *et al.*, 2019; WALKER *et al.*, 1997), propondo diferentes modelos para avaliação humana de *chatbots*.

Portanto, uma vez que essas estruturas dependem da avaliação humana, pode-se dizer que não são facilmente escaláveis. Em (PRZEGALINSKA *et al.*, 2019) os autores argumentam que a confiança está no centro da interação entre um *chatbot* e um humano e analisam como a confiança, sendo extremamente importante, está sendo redefinida com a introdução de assistentes virtuais aptos para aprendizado profundo. No entanto, a métrica de avaliação proposta parece não considerar eficiência, coesão e qualidade geral do diálogo.

Métricas de avaliação automatizadas são mais eficientes em termos de tempo e recursos necessários para realizar a avaliação. No entanto, ainda parece haver uma falta de padrões de indústria em termos de métricas de avaliação aplicadas, e parece faltar nas métricas de avaliação automatizadas a capacidade de avaliar corretamente a qualidade, eficiência e eficácia da conversa como um todo. Porém, dado o fato de que essas métricas podem ser mais facilmente utilizadas, eles ainda são amplamente implementados para avaliar *chatbots*. As métricas de avaliação usadas para medir a precisão serão métricas de avaliação padrão usadas para tradução automática e outras tarefas de Processamento de Linguagem Natural, como BLEU, METEOR e TER, pois, foi usado por (SAIKH *et al.*, 2017; SORDONI *et al.*, 2015). Embora essas métricas

de avaliação sejam consideradas mais adequadas para problemas de tradução automática, eles ainda podem fornecer informações valiosas sobre a implicação textual da saída do *chatbot* (SAIKH *et al.*, 2017).

O F-score, alternativamente referido como F1-score, é uma estatística que indica o quão preciso é um modelo aplicado a um determinado conjunto de dados. É usado para avaliar sistemas de classificação binária que categorizam avaliações de exemplo como “positivas” ou “negativas”. O F-score é uma medição da precisão e *recall* do modelo; é definida como a média harmônica entre a precisão e *recall* do modelo. A equação é apresentada abaixo:

$$F_1 = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

O F-score é frequentemente usado para avaliar sistemas de recuperação de informação, como motores de busca, bem como vários tipos de modelos de aprendizado de máquina, principalmente em processamento de linguagem natural. O F-score pode ser ajustado para priorizar a precisão acima do *recall*, ou vice-versa. As F0.5- e F2-scores, assim como o F1-score normal, são muitas vezes usados como F-scores ajustados. O F-score de um modelo perfeito é 1 (SASAKI, 2007). Essa métrica de avaliação tem sido aplicada em poucas pesquisas para avaliar a performance dos *chatbots*, como em (CUAYÁHUITL *et al.*, 2019; XU *et al.*, 2020).

A perplexidade, abreviada como PP, é a probabilidade inversa do conjunto de teste normalizada pelo número de palavras. A perplexidade, no entanto, não é aplicável a modelos de linguagem não normalizados (isto é, modelos que não são distribuições de probabilidade reais que somam 1), e é incomparável para modelos de linguagem com vocabulários diferentes (CHEN; BEEFERMAN; ROSENFELD, 1998). Essa métrica, no entanto, tem sido utilizada em alguns estudos recentes para avaliar o desempenho dos *chatbots*, nomeadamente em (DHYANI; KUMAR, 2021; HIGASHINAKA *et al.*, 2014; JOHN *et al.*, 2017). A perplexidade PP de um modelo de linguagem P_M (*nextwordwhistoryh*) em um conjunto de teste T é calculada com a equação:

$$PP_T(P_M) = \frac{1}{\left(\prod_{i=1}^t PM(\omega_i|\omega_1 \cdots \omega_{i-1})\right)^{\frac{1}{t}}}$$

Bilingual Evaluation Understudy (BLEU) é amplamente utilizada para avaliar várias tarefas de PLN, mesmo que tenha sido implementado pela primeira vez para medir as saídas de tradução automática. A métrica BLEU atribui um valor a uma tradução em uma escala de 0 a 1, mas normalmente é expressa como uma porcentagem. Quanto mais próxima a tradução estiver de 1, mais se assemelhará a uma tradução humana. Simplificando, a métrica BLEU conta o número de palavras que se sobrepõem em uma tradução quando comparada a uma tradução de referência, dando às palavras sequenciais uma maior pontuação (Plataforma de tradução automática baseada em KantanMT-Cloud). Os autores (DHYANI; KUMAR, 2021; SAIKH *et al.*, 2017; SORDONI *et al.*, 2015; VASWANI *et al.*, 2017) usaram as pontuações do BLEU para avaliar *chatbots* e outras tarefas de PLN. No entanto, BLEU apresenta alguns problemas. A penalidade de brevidade fixa do BLEU não compensa efetivamente a ausência de recall. Além disso, N-gramas de ordem superior são empregados no BLEU como um *proxy* para o nível de qualidade gramatical de uma tradução. Argumenta-se que uma medida explícita de gramaticalidade (ou ordem das palavras) pode ajudar no nível da qualidade gramatical na tradução pela máquina e resultar em uma associação mais forte com julgamentos humanos da qualidade da tradução. Por fim, o BLEU apresenta uma explicação explícita inadequada na Correspondência de Palavras entre Tradução e Referência; embora as contagens de N-gramas não exijam uma correspondência específica palavra por palavra, isso pode resultar em “correspondências” imprecisas, especialmente para termos de função comuns. Para calcular a pontuação BLEU, primeiro, a média geométrica de precisões de n-grama modificadas P_n é calculada usando n -gramas até o comprimento N e pesos positivos ω_n somando um. c é o comprimento da tradução candidata e r é a referência do comprimento do corpo. Então, a penalidade de brevidade BP é calculada com a equação:

$$BP = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases}$$

Então, BLEU é calculado com a equação:

$$BLEU = BP \cdot \exp \left(\sum_{n=1}^N \omega_n \log P_n \right)$$

Foi criada a *Metric for Evaluation of Translation with Explicit Ordering* (METEOR), em português Métrica de Avaliação de Tradução com Pedido Explícito, expressamente para superar os problemas acima mencionados no BLEU. Ele pontua as traduções com base em correspondências explícitas palavra por palavra entre a tradução e uma tradução de referência. Se muitas traduções de referência estão disponíveis, a tradução fornecida é avaliada independentemente de cada referência e a melhor pontuação é relatada. Isso é explorado com mais detalhes na seção seguinte. METEOR produz um alinhamento entre duas *strings* quando recebe um par de traduções para comparar (uma tradução de sistema e uma tradução de referência). O alinhamento é definido como um mapeamento entre unigramas em que cada unigrama em uma *string* corresponde a zero ou a um unigrama na outra *string* e para nenhum unigrama na mesma *string*. Desta forma, um único unigrama em uma *string* não pode traduzir para mais de um unigrama na outra *string* dentro de um determinado alinhamento (AGARWAL; LAVIE, 2008; BANERJEE; LAVIE, 2005). É usado por (SORDONI *et al.*, 2015) e por (XU *et al.*, 2015) juntamente com o BLEU para avaliar um modelo de *chatbot* e um modelo Legendagem de Imagem, respectivamente. Para calcular a pontuação METEOR para a tradução de uma frase, primeiro, a precisão de unigrama P e unigrama *recall* r são calculados como $P = m / t$ e $R = m / r$, respectivamente, onde m é o número de unigramas mapeados encontrados entre as duas *strings* (uma frase e sua tradução), t é o número total de unigramas na tradução da frase e r é o número total de unigramas na frase de referência. Então, a média harmônica parametrizada de P e R é calculada pela equação:

$$F_{mean} = \frac{P \cdot R}{\alpha \cdot P + (1 - \alpha) \cdot R}$$

A taxa de erro de tradução (TER) tem sido menos utilizada em comparação com outro método para avaliar o desempenho de *chatbots*, mas é amplamente utilizada para avaliar a vinculação textual. TER é uma estatística de avaliação de tradução automática que é calculada

automaticamente. É determinado pela distância de edição. Ele calcula a taxa de erro calculando o número de revisões necessárias para converter uma frase de saída traduzida por máquina em uma frase de referência por tradução humana. Assim, o complemento desta taxa de erro é considerado no cálculo da pontuação de similaridade (DHYANI; KUMAR, 2021; SNOVER *et al.*, 2006). As fórmulas para a pontuação *TER* do computador são apresentadas em equação:

$$TER = \frac{\text{número de edição}}{\text{número médio de palavras referência}}$$

No entanto, todos esses modelos de avaliação baseados em n-grama parecem ser menos adequados para avaliar sistemas de diálogo em comparação com outras tarefas de PLN, porque duas respostas podem não ter sobreposição de n-gramas, mas eles podem ser igualmente eficazes em responder a uma determinada mensagem. Por esta razão, alguns trabalhos recentes (KANNAN; VINYALS, 2017) foram conduzidos para estudar o uso de um método de avaliação adversarial para avaliar modelos de diálogo. Considerando a eficácia de redes generativas adversariais (GANs) para geração de imagens, os autores propõem que um indicador da qualidade de um modelo é a facilidade com que sua saída pode ser distinguida da de uma pessoa. Para começar, eles pegam um modelo implantado de conversa totalmente treinado em escala de produção como parte do sistema *Smart Reply* (o "gerador") e treinam um segundo RNN (o "discriminador") na seguinte tarefa: dada uma mensagem recebida e uma resposta, deve prever se a resposta foi amostrada do gerador ou de um humano.

O objetivo desta seção é determinar se um arranjo contraditório é viável para avaliação. Os autores podem demonstrar que um discriminador pode separar com sucesso a saída do modelo e a saída da produção humana em mais de 60% das vezes. Além disso, parece expor as duas principais falhas do sistema: uma distribuição de comprimento errônea e uma dependência em respostas típicas e básicas, como "Obrigado".

No entanto, grandes dificuldades com a aplicação real desta estratégia persistem. Ainda não há indicação de que um modelo com menor acurácia do discriminador é necessariamente superior na avaliação humana. No entanto, a abordagem parece interessante, pois reproduz essencialmente um teste de Turing em um modelo automatizado e escalável. (KANNAN; VINYALS, 2017; LI *et al.*, 2017) sugerem o emprego de treinamento adversário para geração de diálogos em domínio aberto, inspirando-se no teste de Turing: o sistema é treinado para gerar sequências que são indistinguíveis de sentenças geradas por humanos.

Juntamente com treinamento contraditório, eles descrevem um modelo para avaliação adversária que aumenta o sucesso em enganar um adversário como critério para avaliar os diálogos, evitando uma série de perigos potenciais. Os autores (GHANDEHARIOUN *et al.*, 2019; KUKSENOK; MARTYNIV, 2019) propõem diferentes estruturas de avaliação. Apesar disso, as estruturas propostas parecem ser inadequadas em domínio aberto, *chatbots* generativos, e não foram exaustivamente testados.

4.5. Aplicação de chatbos

Os *chatbots* são aplicados em muitos domínios diferentes. No que diz respeito à Educação e à Pesquisa, parecem ser principalmente baseados em recuperação de informações ou AIML. Pouco a nenhum aplicativo de *Deep Learning* foi usado nesses campos. A escolha parece justificada pelo fato de que os *chatbots* criados para fins educacionais muitas vezes visam fornecer informações específicas (como horários de aula) ou material educacional, conforme demonstrado pelos autores (ALASUNDRAM *et al.*, 2019; ARIFI; EBNER; EBNER, 2019; AUGELLO *et al.*, 2012; BALA *et al.*, 2017; BERGER; EBNER; EBNER, 2019; MIRHET; EBNER; EBNER, 2020; NWANKWO, 2018; WANG; PETRINA, 2013), que fornecem exemplos de *chatbots* aplicados à Educação e Pesquisa. Por motivos semelhantes, como no campo da educação, a maioria orientados para cuidados de saúde são baseados em Recuperação de Informações. Os autores (ATHOTA, 2020; AYANOUZ; ABDELHAKIM; BENHMED, 2020) fornecem diferentes exemplos de aplicações de assistentes virtuais em cuidados de saúde.

Para os autores (CUI *et al.*, 2017; OIKUMORO; JAWAD, 2019; SINGH *et al.*, 2018), os *chatbots* específicos para *e-commerce* apresentam diferentes configurações, principalmente configurações baseadas em recuperação de informações, exceto com alguns algoritmos de *Deep Learning* também envolvidos na arquitetura geral. Isso ocorre possivelmente porque no *e-commerce*, eles são frequentemente usados para fornecer suporte ao cliente. Portanto, devem ser capazes não apenas de fornecer informações sobre o catálogo de produtos e experiência de compra, mas também conversas com o consumidor. Nesses casos, os autores fornecem diferentes exemplos de assistentes virtuais aplicados ao comércio eletrônico.

Outros aplicativos de *chatbots* baseados em recuperação de informações podem ser encontrados em Treinamento (CASILLO *et al.*, 2020), Tecnologia da Informação (MELO *et al.*, 2020) e Finanças (OKUDA; SHODA, 2018), possivelmente por motivos semelhantes.

Os modelos sequência-a-sequência com mecanismos de atenção têm sido utilizados para prestar serviços de Direito e Recursos Humanos (JOHN *et al.*, 2017;

SHEIKH; TIWARI; SINGHAL, 2019). No primeiro caso, ele é baseado em Long Short-Term Memory (LSTM), enquanto no segundo caso é baseado em Unidades Recorrentes com Porta Bidirecional. Como mostrado, quando aplicado a domínios específicos os *chatbots* tendem a recorrer a sistemas de recuperação de informações e sistemas baseados em regras, ou uma combinação dos dois. Apenas alguns exemplos de aplicativos usaram tecnologias de *Machine Learning*. Isso pode ser devido a fatores como:

- *Machine Learning* em geral e *Deep Learning* em particular, exigem uma grande quantidade de dados de treinamento; embora os dados de treinamento estejam se tornando cada vez mais disponíveis, encontrar um conjunto de dados adequado ainda pode representar um desafio. Além disso, os dados precisam ser pré-processados para serem usados e podem conter ruídos indesejados.
- O treinamento é caro em termos de infraestrutura e recursos humanos, e consome muito tempo também.
- Os *chatbots*, quando não são usados para fins sociais ou complementares, geralmente são aplicados para um domínio específico, o que significa que eles exigem dados de treinamento específicos do domínio (por exemplo, informações e detalhes de produtos, informações financeiras, material educacional, informações de saúde). Este tipo de dado é muitas vezes confidencial devido à sua natureza; elas não estão prontamente disponíveis em domínio público para treinar um mecanismo de *Deep Learning*. Além disso, dada a natureza dos dados necessários e das tarefas, o *chatbot* é necessário realizar (por exemplo, acessar o histórico de compras de um cliente ou fornecer mais informações sobre um recurso do produto), e, portanto, a Recuperação de Informações pode ser a melhor solução para a maioria das aplicações de casos de uso.

5. DESAFIOS E LIMITAÇÕES NO DESENVOLVIMENTO DOS CHATBOTS

No presente capítulo são apresentados os desafios e limitações ligados ao desenvolvimento dos softwares de diálogo inteligente.

Conforme apresenta (CALDARINI; JAF; MCGARRY, 2022), apesar dos avanços atuais nas áreas de *Deep Learning* e PLN, as arquiteturas dos *chatbots* ainda apresentam algumas deficiências. Em primeiro lugar, os diferentes modelos de linguagem propostos como arquitetura de *chatbots* ainda não conseguem imitar a conversa humana devido à abordagem incorreta da modelagem de diálogo. O problema subjacente é que este modelo tenta resolver

problemas de conversação com uma abordagem “próximo passo”: dada uma entrada, ele tenta prever a melhor saída. Isso, porém, não é o raciocínio por trás da conversa humana, que não avança simplesmente um passo de cada vez tempo, mas sim levando em consideração uma série de etapas anteriores, como contexto da conversa e as informações que estão sendo compartilhadas entre os participantes (VINYALS; LE, 2015).

A conversação humana não é um processo passo a passo como é modelado nas arquiteturas dos *chatbots*, mas sim uma jornada contínua, um vai-e-vem contínuo, onde cada passo é dependente dos anteriores ou posteriores. Essa dependência constitui o contexto conversacional, e mesmo que alguns novos modelos tenham tentado capturar tal contexto, progressos ainda devem ser feitos (ADIWARDANA *et al.*, 2020; CHRISTENSEN *et al.*, 2018; DAI *et al.*, 2019).

Outra falha bastante importante na arquitetura é a aparente falta de um modelo de IA para assistentes virtuais de recuperação de informações. Esses, como evidenciado na literatura sobre aplicativos de *chatbots*, são amplamente populares em todos os setores (por exemplo, saúde, comércio eletrônico, atendimento ao cliente e educação) porque são capazes de fornecer respostas coerentes a um tópico específico (por exemplo, agendar uma consulta, especificações do produto, devolver um item e encontrar materiais de estudo), uma vez que podem encontrar uma resposta semelhante em sua base de conhecimento (CALDARINI; JAF; MCGARRY, 2022).

Atualmente, é possível perceber os modelos de *chatbots* de recuperação de informações dependem do conjunto de dados usado para treiná-los e não há um modelo de aprendizado flexível que pode ser aplicado a diferentes conjuntos de dados. Parece, de fato, que a pesquisa agora é mais focada em grandes modelos generativos em vez de menores, facilmente implementados e independente de domínio. Esse modelo encontraria várias aplicações em todos os setores. O desafio no desenvolvimento de tais modelos é a falta de um domínio específico de acesso aberto, dados linguísticos, bem como a natureza altamente diversificada de indústrias e tópicos onde tais modelos seriam aplicados (CALDARINI; JAF; MCGARRY, 2022).

Em termos de aplicações ainda há uma lacuna a ser preenchida entre os modelos utilizados pela indústria e as recentes melhorias na área. Os grandes modelos que representam a mais recente melhoria no campo, sejam eles modelos de linguagem em geral ou modelos de diálogo, ainda não são adequados para serem facilmente aplicados pela indústria, pois requerem grande poder computacional e conjuntos de dados de treinamento excepcionalmente grandes.

Como afirma (CALDARINI; JAF; MCGARRY, 2022), os *chatbots* são aplicados em diferentes indústrias para atender a necessidades específicas.

Para cumprir seu propósito, os modelos dos assistentes virtuais são treinados em conjuntos de dados específicos e suas aplicações dependem de frameworks diferentes, muitas vezes complexos, que incluem gerenciadores de diálogo e/ou árvores de decisão, e são baseados em conhecimento ou em regras. No entanto, dado o uso generalizado de tais modelos, parece evidente que soluções sob medida que garantem respostas qualitativas e precisas para consultas de clientes específicos são preferidas em modelos maiores que exigem uma grande quantidade de dados e um melhor desempenho em conversas de domínio aberto, mas pode não ter um desempenho tão bom em um domínio fechado. É evidente que há uma divisão entre domínio aberto (orientado à pesquisa modelos) e domínio fechado (aplicações orientadas para a indústria). Para preencher esta lacuna, modelos flexíveis, menores e menos dependentes de domínio seriam melhores. No entanto, desenvolver estes modelos ainda é desafiador porque *chatbots* generativos de domínio aberto tendem a ser menos precisos e eficientes em suas respostas (CALDARINI; JAF; MCGARRY, 2022).

Em relação à avaliação do *chatbot*, existem pelo menos duas grandes limitações que ainda precisam ser abordadas. Em primeiro lugar, não existe uma estrutura comum para a avaliação. Embora algumas métricas sejam amplamente utilizadas para medir o desempenho, não há métricas específicas ou conjunto de métricas comumente usadas como referência. A falta de um quadro comum de referência sobre a avaliação limita o teste e a comparação correta de diferentes modelos. Essa limitação também pode ser devido à segunda limitação que surge da literatura: a falta de um método de avaliação automática confiável e eficiente (CALDARINI; JAF; MCGARRY, 2022).

No entanto, a avaliação não é aprofundada em trabalhos recentes e, embora tenha sido apresentado algumas peças relevantes da literatura sobre avaliação, como foi discutido, elas se concentram mais em estruturas únicas de avaliação em vez de comparar várias métricas de avaliação. Tal análise foi fornecida, comparando-se diferentes métricas de avaliação e discutiu-se outras etapas a serem tomadas para melhorar a avaliação dos *chatbots*.

Conforme destaca (CALDARINI; JAF; MCGARRY, 2022), os métodos de avaliação automática apresentam algumas deficiências, e nenhum é capaz de avaliar completamente aspectos importantes da qualidade do diálogo, como aspectos de coerência, coesão e vinculação textual. Por esta razão, muitos modelos dependem da avaliação humana, mas a avaliação humana é cara, demorada, não facilmente escalável, sujeita a vieses e apresenta falta de

coerência. Além disso, nem mesmo a avaliação humana apresenta um quadro de referência comum entre os modelos. Para superar essas limitações, um novo método de avaliação automática confiável deve ser proposto. Tal método deve oferecer resultados de conversação de *chatbots* e estimativa qualitativa com base na correção, coerência e coesão do diálogo em vários contextos. Tal método de avaliação pode representar um passo crucial para melhorar o desempenho geral.

Algumas lacunas de pesquisa apontadas por (CALDARINI; JAF; MCGARRY, 2022) sobre os assistentes virtuais que precisam ser destacadas. Em primeiro lugar, embora muitos artigos de pesquisa sobre *chatbots* apresentem uma explicação detalhada das tecnologias e implementações, pesquisas recentes carecem de informações sobre a maioria dos recentes avanços em modelos de linguagem que podem ser aplicados, como *Transformers*, onde uma visão geral desses módulos avançados foi fornecida. Entretanto, uma análise aprofundada desses modelos e sua aplicação em *chatbots* seria bastante útil. Da mesma forma, poucas informações e análises sobre conjuntos de dados são fornecidas. O tipo e a qualidade dos dados usados para treinar modelos de *Deep Learning* é particularmente importante para determinar a saída e a precisão do modelo. Isso é particularmente verdadeiro em modelos de linguagem, uma vez que o modelo tem que aprender a tarefa de linguagem com base nos dados linguísticos disponíveis, e interpretar dados linguísticos não é tão simples quanto interpretar dados numéricos. Por essas e outras razões, é importante discutir e analisar os dados usados para treinar os diferentes modelos. Tal análise também é importante porque permite uma comparação mais justa de diferentes modelos e sua eficiência.

Uma análise dos conjuntos de dados de domínio aberto mais comuns é feita no trabalho de (CALDARINI; JAF; MCGARRY, 2022). Finalmente, um aspecto vital dos *chatbots* é sua aplicação em diferentes indústrias e em cenários da vida real. Mesmo que alguns artigos forneçam uma visão geral da aplicação e alguns números sobre o seu aumento em popularidade na web, parece haver uma falta de clareza na identificação da indústria ou aplicativos orientados a tarefas e *chatbots* sociais ou complementares. Além disso, esta distinção não é feita ao discutir diferentes modelos e sua aplicação.

A modelagem é uma tarefa interessante que combina *Deep Learning* e PLN, e cujas aplicações cresceram incrivelmente nos últimos anos. Embora os primeiros *chatbots* tenham surgido há sessenta anos, o campo continuou se expandindo e apresenta desafios novos e envolventes. Sua implementação em vários setores e como assistentes cria muitas

oportunidades e fascinantes caminhos de pesquisa, como mostra a notável quantidade de literatura publicada sobre o assunto nos últimos anos.

Leal (2018) mostra também que outra limitação são os falsos positivos que ocorrem quando um *chatbot* responde incorretamente a uma pergunta ou comando. Isso pode ser por vários motivos, incluindo:

- 1- Compreensão limitada da linguagem natural: embora os chatbots sejam projetados para entender a linguagem humana, eles ainda têm limitações e podem interpretar mal perguntas ou comandos;
- 2- Dados de treinamento inadequados: se um *chatbot* não for treinado adequadamente com uma variedade suficiente de perguntas e respostas de amostra, ele poderá responder incorretamente a novas situações.
- 3- Problemas de acessibilidade: Se um usuário tiver uma deficiência ou não falar o idioma nativo do *chatbot*, pode haver uma barreira para entender e responder adequadamente.

Para minimizar os falsos positivos, é importante treinar os chatbots com uma ampla variedade de dados e prestar atenção a qualquer padrão de comportamento incorreto. Além disso, é útil fornecer um mecanismo para que os usuários relatem bugs e forneçam feedback aos desenvolvedores para melhorar continuamente a precisão do chatbot (LEAL, 2018).

Por fim, deve ser observado que diferentes níveis de incorporação - a maneira como os *chatbots* se assemelham aos humanos (SUNDAR; GO, 2019) - e a divulgação - como e quando a natureza do *chatbot* é revelada ao usuário - tem significativo impacto o engajamento e a confiança dos usuários (LUO *et al.*, 2019) que usam as ferramentas de diálogo automatizado.

Para concluir o capítulo, um outro ponto importante a ser discutido é o assédio que acontece diversas vezes durante as interações com os assistentes virtuais, e em especial ao *chatbot* BIA do banco Bradesco. Desde o início, quando a BIA passou a atender clientes e não somente funcionários internos, episódios indesejados e de caráter ofensivo foram relatados pelo banco, o que evidencia que mulheres são as principais vítimas de assédios sexuais, morais e sofrem importunação sexual. Mesmo não sendo um pessoa real, muito menos do sexo feminino, a BIA contém elementos femininos em sua implementação e forma de interagir que confirma que esse tipo de ataque acontece por conta do gênero (Bradesco, 2023).

Para entendermos brevemente sobre assédio, podemos considerar os tipos de assédio e suas definições. Assédio virtual é “a prática de comportamentos hostis, ofensivos e violentos contra uma pessoa ou um grupo através de tecnologias e plataformas online” (Bradesco, 2023).

Já assédio moral é definido pelo (Bradesco, 2023) como:

(...) a humilhação ou ofensa a uma pessoa ou um grupo que acontece diversas vezes no ambiente de trabalho. Quando esses ataques acontecem de forma isolada, chamamos de violência psicológica. O assédio moral pode se apresentar de formas variadas, como discriminação racial, homofobia, transfobia, discriminação por algum tipo de deficiência e discriminação linguística.

Ainda pelo (Bradesco, 2023), a empresa define assédio sexual da seguinte forma:

(...) é o ato de constranger alguém para conseguir favorecimento sexual, baseado em uma relação de hierarquia e subordinação entre vítima e agressor. O assédio sexual pode acontecer uma única vez - diferente do assédio moral - e é restrito ao ambiente de trabalho. Recentemente, a importunação sexual foi tipificada em Lei e considera casos que não envolvem uma relação hierárquica ou de trabalho.

Pode-se observar que mesmo não se tratando de uma pessoa real, esse tipo de problema acontece principalmente quando se trata de pessoas do gênero feminino, e não falar sobre e combater esses casos pode contribuir para a naturalização desses comportamentos. O Bradesco mudou as respostas da BIA quando mensagens de assédio são enviadas e não mais toleram esse tipo de interação, além de reforçar o compromisso de lutar contra essas formas de violência através do movimento “Hey, atualize minha voz” (Bradesco, 2023).

O autor de (NeoAssist, 2020), reforça que o viés inconsciente tem um papel fundamental no que diz respeito a esses tipos de assédio, uma vez que ele molda a forma como uma sociedade se comporta. A ligação desse tema nos mostra que, além de programar os *chatbots* para que respondam bem à essas violências e inibindo esses comportamentos, é importante também investir em campanhas de conscientização e educação e desconstruir aos poucos esse cenário.

6. ESTUDO DE CASO: BIA – BRADESCO INTELIGÊNCIA ARTIFICIAL

No presente capítulo é feito um estudo de caso da ferramenta BIA (BRADESCO, 2022). Inicialmente é relatado um histórico do surgimento da ferramenta. Em seguida, são destacados

detalhes de implementação. Na sequência, são apresentadas simulações de uso e uma análise das interações dos clientes do banco com o *chatbot*. Por fim, é discutido o impacto da ferramenta no setor bancário.

6.1. História do surgimento da BIA

Segundo (PAIVA, 2018), a criação de uma assistente virtual pode levar muito mais tempo que a gestão de um ser humano, devido à complexidade de sua implementação e toda as outras vertentes discutidas neste trabalho. O autor afirma ainda que por se tratar de serviços no âmbito financeiro, a situação se torna ainda mais difícil e demorada, já que se deve cumprir uma série de regulamentações e abordar questões de segurança dos dados. Por este motivo, a “gestação” de um *chatbot* como este pode levar anos para ser concluída. Destaca ainda que o Bradesco optou por escolher um bom líder para cuidar de toda a parte da pesquisa sobre a área, sendo um profissional que entende bem do assunto inteligência artificial e segurança bancária. O nome escolhido foi o de Marcelo Câmara, sendo conhecido como um dos pais da BIA (BRADESCO, 2022), juntamente com toda a tecnologia Watson da mundialmente conhecida IBM.

Estima-se que essa gestão tenha durado por volta de dois anos e nesse primeiro momento seu uso foi restrito aos funcionários do banco, e posteriormente liberada para os clientes do Bradesco. Na época de 2018, a BIA (BRADESCO, 2022) apenas conseguia tirar dúvidas dos usuários, porém, com a promessa de que conseguiria futuramente realizar transações bancárias. Destaque para a implementação do sistema também nos aplicativos de mensagens Messenger e WhatsApp.

Câmara relembra das dificuldades e dos principais desafios encontrados durante o processo. Ele diz que apesar do contato com a IBM ter acontecido muito tempo antes ao ano que de fato a parceria foi fechada, ele sempre trabalhou com o objetivo de estar à frente nas pesquisas e inovação de produtos, e que por este motivo, de olho em tendências, pode fechar um acordo comercial para que fechassem um contrato em 2015, classificando o projeto como secreto e que em maio de 2018 o trabalho completou três anos.

Paiva (2018) ainda diz que inicialmente foram fornecidos dados na língua portuguesa para a Watson, e que suas primeiras interações aconteceram com os próprios funcionários do banco. A BIA (BRADESCO, 2022) se tornou um sucesso logo no começo, respondendo dúvidas dos gerentes que antes era atendidas por um departamento de *call center* interno de

forma mais rápida, tornado o atendimento desses funcionários mais eficiente, já que dependia de uma solução imediata para as questões uma vez que os clientes estavam frente a frente com os clientes. Mesmo com uma interface simples a interação era efetiva e tudo que era comunicado através do sistema servia para gerar um grande conjunto de dados para expandir o vocabulário e tornar o *chatbot* mais inteligente. Pouco depois, de acordo com a ampliação do uso para outras centenas de agência, o sistema foi se tornando cada vez melhor e mais rico, possibilitando até o treinamento de funcionários do banco HSBC, que em 2016 foi comprado pelo Bradesco.

Os grandes benefícios citados já mostraram que apesar do custo do desenvolvimento da tecnologia e sua implantação, o investimento acabou se pagando, possibilitando a criação de novos recursos e atingir mais produtos nos atendimentos. “A ideia era chegar a 11 produtos e serviços até setembro de 2016, mas atingimos isso em março daquele ano. E em setembro de 2016 estávamos com 59 serviços compreendidos pela BIA (BRADESCO, 2022)”, afirma (PAIVA, 2018).

Em 2017, 16 milhões divididos entre clientes e funcionários já o *chatbot*. A grande maioria desse número, que são clientes, foi introduzida logo após o contato inicial por parte dos funcionários das agências e em alguns pontos de atendimento. Percebeu-se que estava pronta para que os clientes também a utilizassem para os atendimentos e inicialmente foi implantada para um grupo seletivo de clientes, denominados clientes prime. Isso ajudou também a verificar se a tecnologia suportava esse número gigantesco de usuários para logo depois ampliar para o restante dos clientes de outros segmentos.

Algumas transações puderam ser implementadas também além de apenas responder os clientes. Exemplos como pedir para a BIA (BRADESCO, 2022) fazer uma transferência nos leva para a tela onde o cliente preenche os dados da conta destino, valor e finaliza a transação. O objetivo também era que além de levar para a página solicitada, a IA também realizasse a transação. Isso nos leva ao fato de que a IBM Watson possuía o objetivo de entender o que o usuário está querendo fazer, sendo indiferente a forma e as palavras que ele utilizou. Interações usando palavras regionais, palavrões, abreviações e formas diferente de dizer a mesma coisa foram feitas a fim de testar se a BIA (BRADESCO, 2022) de fato conseguia entender a intenção do usuário, e o resultado foi satisfatório. Muito trabalho tem sido feito para que o contexto das interações seja bem interpretado, já que humanos possuem essa capacidade e a intenção da IA é ser cada vez mais semelhante ao homem. Vale destacar também que a linguagem informal é importante, uma vez que se assemelha a utilizada pela maioria dos usuários. “Um exemplo seria

escrever “obg” no lugar de “obrigadao”. É entendido que é essas palavras ou gírias ou abreviações e são aceitáveis naquele contexto “, afirma (PAIVA, 2018).

6.2. Implementação da BIA: Pontos de Destaque

Conforme abordado pelos autores, do ponto de vista funcional, percebe-se que a IA Banco Bradesco responde de forma passo-a-passo, na medida que uma ação só ocorre em decorrência de uma primeira ação. Apresentando assim ser um *chatbot* baseado em regras. Apesar de responder que ainda está aprendendo conforme as interações humanas vão acontecendo, pode-se perceber que seu comportamento se dá através de pesquisas e respostas comuns. Na tela inicial dentro do aplicativo, quando se acessa a BIA (BRADESCO, 2022), o *chatbot* oferece algumas opções pré-programadas e o que não se encontra dentro desses contextos parece não conseguir uma resposta satisfatória.

No quesito personalidade, foi adotada uma personalidade neutra. Mas imagina-se que no futuro será possível variar a personalidade do *chatbot* de acordo com o cliente e com o contexto das entradas apresentadas.

Um ponto interessante fica acerca do processamento dos dados. O processamento do reconhecimento da intenção é feito pela IBM. A transcrição de fala para texto é feita no próprio dispositivo do cliente e às vezes vai para a nuvem também (PAIVA, 2018).

O conjunto de dados fornecido para a BIA (BRADESCO, 2022), assim como ela responde no chat, vem das inúmeras conversas e interações com os usuários e isso acaba limitando a coerência e eficiência dos resultados. O que pode parecer uma pergunta simples se torna uma dificuldade para o *chatbot* buscar em seu banco de dados algo que retorne um verdadeiro positivo para o que está sendo pedido, justamente pelos gargalos apresentados no decorrer deste trabalho, como conjunto de dados limitados e questões éticas relacionadas, por exemplo, a dados privados dos clientes de suas contas bancárias.

Após algumas interações basicamente de perguntas/palavras e respostas, um *feedback* é solicitado ao usuário perguntando se a resposta dada foi satisfatória e se a dúvida foi atendida. Caso selecionado a opção de “não” um menu contendo mais perguntas e respostas pré-programadas é apresentado para tentar ajudar ao usuário a encontrar o que procura, deixando assim a própria IA em segundo plano, e se tornando apenas uma página de Perguntas Frequentes.

Ainda no contexto do *feedback*, o fato que o torna um dos pontos mais importantes acerca de sua implementação se deve ao aprendizado que a máquina tem ao receber afirmações ou negativas a cada resposta dada. Assim, a IA consegue aprender conforme as interações vão acontecendo e os usuários a "ensinam" que o que foi abordado na comunicação foi assertiva ou não ajudou. Se a resposta dada pela BIA (BRADESCO, 2022) para o usuário atendeu as expectativas e o feedback for positivo, o *chatbot* então entende que assim que outra pessoa usar palavras-chaves similares, portanto, a dúvida for parecida, haverá fortes indícios de que a IA irá devolver a mesma resposta. O feedback é dado através dos botões representados por um joinha para positivo e joinha para baixo para negativo, conforme exibido no fim da Figura 3.

À medida que a limitação dos conjuntos de dados necessários para treinar o *chatbot* se torna um problema, o impacto na entrega de resultados e na avaliação da eficiência, coesão e qualidade geral do diálogo se torna grande e prejudicial na experiência como um todo. Um modelo de aprendizagem para contextos específicos como esse ainda precisa ser estudado e novas formas de obter esses dados, assim como estruturas de avaliação eficientes, devem ser analisados e explorados em futuras pesquisas, pois atualmente, conforme foi apresentado, não há um modelo que atenda aos requisitos necessários para se ter uma experiência mais próxima a de uma interação com humanos.

No quesito segurança, não fazem ainda o tráfego de dados sensíveis. Está sendo estudado a melhor maneira de fazê-lo. Tem sido estudado modelos de mascaramento, para que se o dado for para fora, que seja de maneira mascarada. Ainda não se há um modelo fechado. É um desafio, mas está sendo equacionado. Sabe-se o caminho e faltam poucos detalhes. A ideia é que o aplicativo funcione de forma bem segura, onde acredita-se que mascarando já se resolve muitos casos. O computador que vai entender a intenção não precisa saber agência e conta. Dando informações incompletas acredita-se que podem chegar no mesmo resultado (PAIVA, 2018).

Inicialmente as transações serão nos apps. Fora deles seria necessário prover mecanismos de autenticação acopláveis a um WhatsApp, por exemplo, o que até o momento não está implementado. A empresa estima em algum momento uma parceria através de uma combinação entre as companhias (PAIVA, 2018).

No tocante aos benefícios, a agilidade e a qualidade do atendimento são os maiores. A BIA (BRADESCO, 2022) tem uma abrangência que não se teria com um atendente só. Geralmente os atendentes são especializados em algum assunto. A inteligência artificial ainda está em um estágio embrionário. O que dá para fazer será muito mais amplo no futuro. O sistema

é muito bom para big data, para trilhar um bom caminho. Se as pessoas perguntam muito sobre fundo de investimento, se torna sabida a necessidade de melhorar as respostas acerca desse assunto (PAIVA, 2018).

Foram criadas várias equipes com diferentes abordagens de inteligência artificial no Bradesco. Existe a equipe core, que faz a estratégia da IA analisando o que tem mais novo e dando a direção, e também foi formada a equipe de desenvolvedores, que acoplam isso aos sistemas, e uma outra equipe totalmente nova, que é de curadoria. É o pessoal que ensina a BIA (BRADESCO, 2022) a dar respostas adequadas. São eles que olham as tendências de perguntas para dar as melhores respostas para os clientes. A inteligência artificial tem que ser feita com sabedoria humana. Não dá para tirar o humano do circuito. Destaca-se ainda que aproveitaram pessoas que já trabalharam em *call center*. Foram treinadas e dada a elas uma função mais nobre, ao em vez de atender clientes diretamente, elas a treinam a atendente virtual (PAIVA, 2018).

Segundo (PAIVA, 2018), a análise de sentimentos durante as interações está em fase experimental. Apesar disso, alguns conseguem ser identificados: raiva, felicidade e tristeza, basicamente, mas ainda não é possível saber se a pessoa está com saudade, por exemplo. Já para ironia ser identificada a questão se torna mais complicada, já que na fala é mais fácil de se perceber que num texto.

Alguns avanços foram feitos e a BIA (BRADESCO, 2022) está integrada ao WhatsApp. Esse é o diferencial: atendimento no WhatsApp com inteligência artificial. A interação começa com o *chatbot* e, se houver necessidade, é transferida para o gerente. Ela entende a intenção 94% das vezes. Somente em 6% das conversas um atendente humano é acionado.

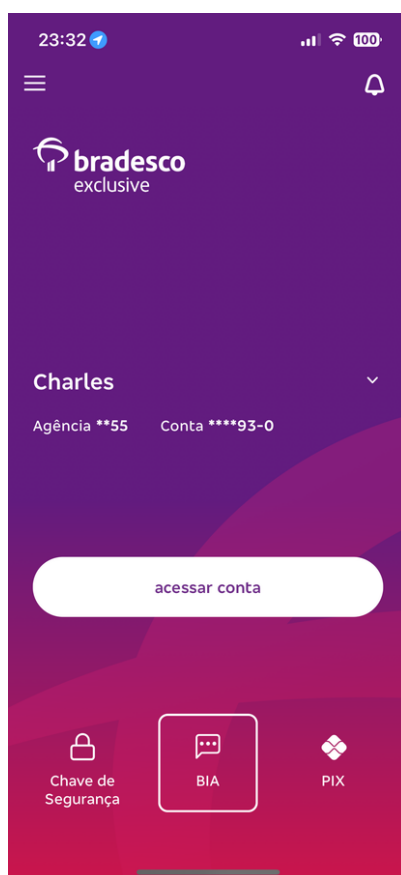
Um outro ponto bastante importante é que o sistema conversacional engloba tudo. A maior parte do que está sendo testado envolve chat. O objetivo final é que ela também atenda o *backoffice* e para outras coisas que não apenas atendimento ou conversação.

6.3. Simulações de uso no aplicativo

Para interagir com o a BIA (BRADESCO, 2022), o cliente deve baixar o aplicativo do Banco Bradesco através das lojas oficiais do sistema operacional do celular. Depois, assim que o app é aberto, o cliente logo é levado para a tela inicial onde existe um atalho para levá-lo direto para o *chatbot*. Nas simulações a seguir deve-se levar em consideração que a conta do cliente não foi acessada. Portanto, informações sobre a conta não serão respondidas se perguntadas à assistente virtual.

Na Figura 1 observa-se a tela inicial do aplicativo e logo no centro inferior observa-se três menus principais: Chave de segurança, BIA e PIX (Instant Payment). O cliente pode acessar a IA tocando no menu “BIA”.

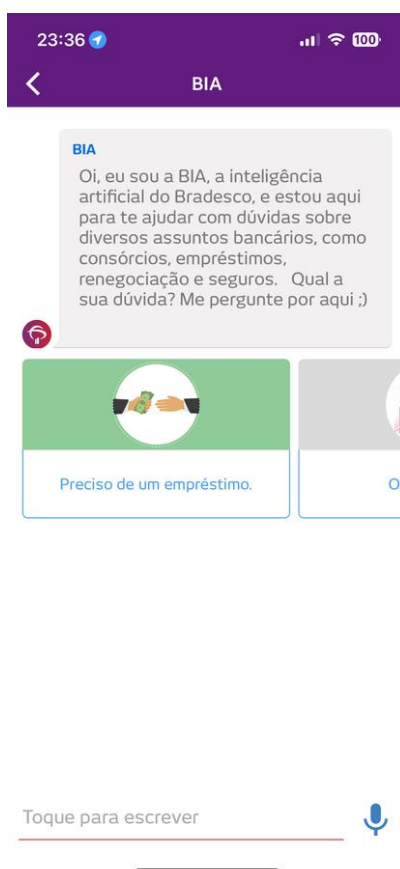
Figura 1 - Tela inicial do aplicativo do Bradesco



Fonte: captura de tela e montagem do autor.

Assim que a BIA é acessada, uma breve introdução sobre o que ela é o que ela pode fazer é exibida. Existem algumas opções pré-programadas para auxiliar o cliente, como a de solicitar um empréstimo, o que é PIX (Instant Payment), como liberar serviços e transações pelos Canais Bradesco e como cadastrar a Chave de Segurança no celular. Porém, se o cliente optar por digitar algo que deseja fazer ou uma dúvida que deseja esclarecer, pode ser utilizado o teclado para escrever ou o microfone para falar o serviço desejado. Na Figura 2 a IA inicia o atendimento dando uma breve introdução do que ela é e como funciona. Também são apresentadas algumas opções mais comuns que são questionadas a ela.

Figura 2 - Tela inicial do menu da BIA

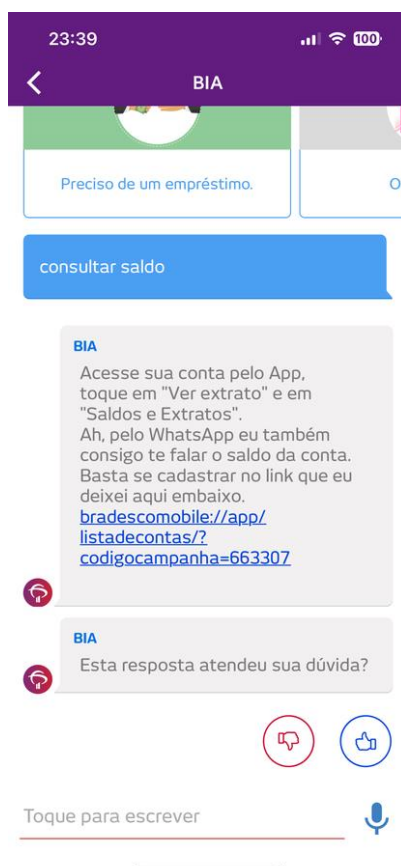


Fonte: captura de tela e montagem do autor.

Na Figura 3 pode ser observado que o texto “consultar saldo” foi digitado e retornou com a resposta de um passo-a-passo de como chegar no valor do saldo em conta. Logo em seguida, ela pede avaliação se a resposta atendeu a dúvida do usuário. Esse *feedback* é extremamente importante para o aprendizado do *chatbot*, como já foi mencionado

anteriormente. Através dele a IA aprende que aquele conjunto de palavras e a resposta dada tem forte conexão e numa próxima interação com outro usuário, por exemplo, a assistente virtual recomendaria o mesmo conteúdo.

Figura 3 - Comando "consultar saldo" sem acesso à conta

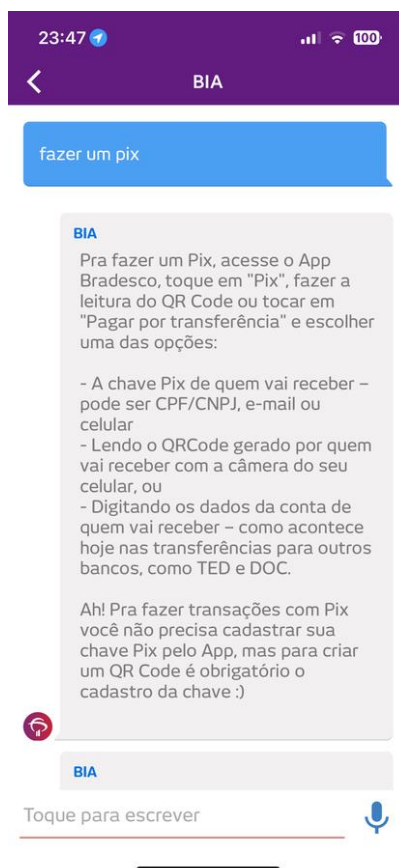


Fonte: captura de tela e montagem do autor.

Como pode-se observar, o comportamento neste caso demonstra que o funcionamento da IA acontece com base em um domínio de regras, como discutido na seção 4.2.1. A IA, através de regras linguísticas e técnicas de correspondência de padrões identifica o que o cliente está pedindo e retorna com uma resposta pré-definida baseada no que foi solicitado.

Na Figura 4 foi enviado o texto “fazer um pix” e a resposta mais uma vez foi satisfatória, uma vez que a BIA responde como executar a transação e uma breve explicação de como preencher os campos e como configurar caso necessário.

Figura 4 - Comando "fazer um pix" e resposta padronizada

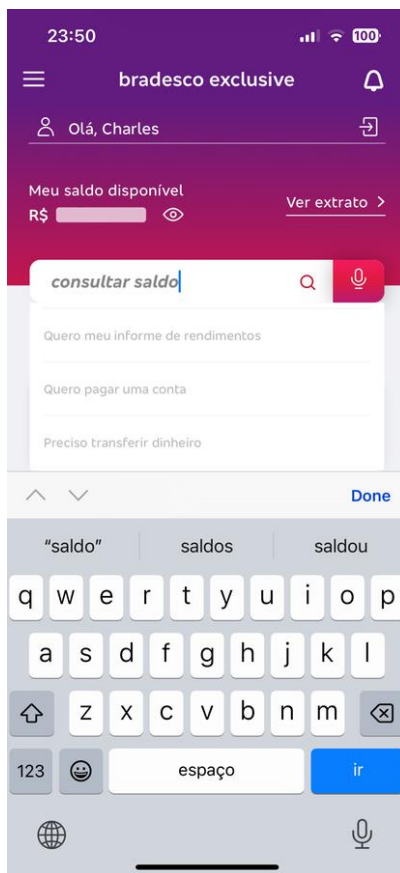


Fonte: captura de tela e montagem do autor.

Estes são alguns exemplos mais simples que mais se parecem como perguntas e respostas. A BIA (BRADESCO, 2022) consegue alterar o comportamento do aplicativo de acordo com o comando enviado, como dito anteriormente, através de texto ou voz.

Mudando o cenário e agora considerando que o cliente acessou sua conta pelo aplicativo, comandos para o *chatbot* podem ser enviados através de uma barra de comandos ou através do ícone do microfone. O mesmo exemplo mostrado anteriormente de consulta de saldo pode ser feito da forma exibida na Figura 5, onde foi digitado o texto “consultar saldo” e depois tocando no ícone da lupa para enviá-lo.

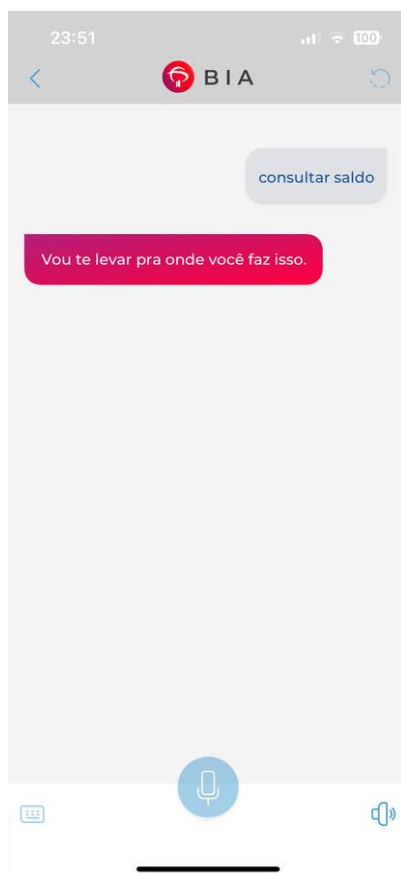
Figura 5 - Comando "consultar saldo" com conta acessada



Fonte: captura de tela e montagem do autor.

Logo depois de enviar o comando, a assistente virtual então é acionada e a solicitação é processada pela inteligência artificial. Na Figura 6 pode-se observar que o *chatbot* leva o cliente para o local mais adequado que corresponde ao que foi pedido, seguindo as regras de coerência as palavras digitadas ou ditadas.

Figura 6 - BIA recebendo o comando e processando a solicitação



Fonte: captura de tela e montagem do autor.

Após um momento rápido, a tela do aplicativo é alterada para corresponder a requisição do cliente, e de modo satisfatório, esse simples comando foi interpretado e atendeu à solicitação exemplificada. Na Figura 7 pode-se concluir que o chatbots conseguiu levar o usuário para a consulta de saldo e os valores da conta do cliente foram ocultados para proteger as informações sensíveis.

Figura 7 - Tela "Consulta de Saldo" redirecionada pela BIA no App do Banco Bradesco



Fonte: captura de tela e montagem do autor.

6.4. Análise das intenções do Bradesco com uso da BIA

Na presente Seção é abordada uma reflexão sobre a análise da campanha que o Bradesco criou para apresentar a BIA (BRADESCO, 2022) e suas intenções. A Tabela 2, apresentada posteriormente, possui os elementos dessa campanha e algumas informações que serviram para conduzir as reflexões sobre o conteúdo mostrado e objetivo a ser atingido pelas estratégias. Assim, conforme a análise desses itens as reflexões obtidas por cada um deles, a tabela foi preenchida com o cruzamento entre as intenções pela autora (SCHORN, 2019) com as hipóteses que serviram de base para essas análises. A autora ainda afirma que existe diversas similaridades e especificidades entre algumas campanhas que o Bradesco lançou para divulgar sua inteligência artificial:

- BIA: a inteligência artificial do Bradesco, lançada em agosto de 2018²;
 - Campanha institucional de lançamento da assistente virtual do Bradesco – BIA.
- Fale com a BIA e abra sua conta, lançada em janeiro de 2019³;
 - Incentivo para abertura de novas contas;
- É você quem inspira a BIA, lançada em abril de 2019⁴.
 - Manifesto da BIA - Campanha que incentiva seu uso para aprimorar cada vez mais seu desempenho.

Essas campanhas e suas intenções compartilham de uma mesma base, que vai da realidade dos clientes e de outros usuários para construir a história da BIA (BRADESCO, 2022) e sua utilidade no cotidiano das pessoas que utilizam os mais diversos serviços bancários. Isso mostra a utilidade do *chatbot* em relação aos casos reais das pessoas em resolver questões de forma simples, disponível a todo momento na palma da mão, através do smartphone, e que também atinge um público maior.

O Bradesco foi a empresa do segmento bancário pioneira no lançamento de uma assistente virtual que usa inteligência artificial e mostrou-se à frente da concorrência quando possibilitou a abertura de contas através do aplicativo de forma intuitiva e com a agente da BIA (BRADESCO, 2022). A facilidade no uso do aplicativo com instruções da assistente virtual inteligente é demonstrada na primeira campanha mencionada, exemplificando alguns comandos de voz e como a IA reage a esses comandos.

O tema da segunda campanha incentiva a abertura de contas destacando os diferenciais do banco. O ponto interessante aqui é que a BIA (BRADESCO, 2022) se torna uma espécie de garota-propaganda, entregando aos usuários e possíveis novos clientes alguns motivos pelos quais o Bradesco se destaca com a concorrência, de forma a apresentar uma maneira inteligente de guiar esse usuário no uso do aplicativo, por exemplo.

Já na terceira campanha a BIA (BRADESCO, 2022) é apresentada com uma voz robótica, com o intuito de simular a inteligência artificial falando direto com o usuário, e reforça

² Disponível em: <<https://www.youtube.com/watch?v=k3brZzuC5Ug>>. Acesso em: 25 Set. 2022.

³ Disponível em: <<https://www.youtube.com/watch?v=Loyq4N6rV5c>>. Acesso em: 25 Set. 2022.

⁴ Disponível em: <<https://www.youtube.com/watch?v=CAAACetLLm0>>. Acesso em: 26 Set. 2022.

a ideia de que conforme as interações entre ela e os humanos (clientes e não clientes) acontece, a assistente virtual evolui e absorve cada vez mais conhecimento.

Por fim, percebe-se que o objetivo geral dessas campanhas é apresentar uma tecnologia inovadora, que coloca o Bradesco numa posição de destaque e que a BIA (BRADESCO, 2022) pode ajudar na rotina de quem utiliza aplicativos bancários, reforçando mais de uma vez que seu uso deve ser feito mesmo que ela ainda não responda corretamente num estágio inicial pois seu aprendizado vem das interações. (SCHORN, 2019)

Tabela 2 - Análise de dados das peças

Campanha	Descrição	Elementos	Intenção	QR CODE
BIA: a inteligência artificial do Bradesco	Campanha institucional de lançamento da BIA. BIA é apresentada como uma facilitadora.	• <i>Storytelling</i> do roteiro; • Textos; • Elementos visuais; • Relação IA e humanos.	Bradesco apresenta a BIA como uma tecnologia inovadora que dá vida à uma nova realidade de relacionamento que já está acessível à população em geral e que tem como objetivo facilitar a rotina financeira das pessoas e o contato entre clientes e o banco.	
Fale com a BIA e abra sua conta	Campanha promocional para abertura de conta. BIA é apresentada como uma facilitadora e uma porta voz dos benefícios do banco.	• <i>Storytelling</i> do roteiro; • Textos; • Elementos visuais; • Relação IA e humanos.	Bradesco utiliza a BIA como porta-voz para apresentar se como um banco que está à frente de seus concorrentes, para destacar seus diferenciais, sendo que um destes é a própria conveniência em ter a BIA como facilitadora da rotina financeira.	

É você quem inspira a BIA	Campanha de manifesto da BIA. BIA como a IA do Bradesco.	• <i>Storytelling</i> do roteiro; • Textos; • Elementos visuais; • Relação IA e humanos.	Bradesco abre o diálogo sobre a relação necessária entre a BIA e os humanos com a intenção de gerar mais interações e empatia com o processo de aprendizado da IA.	
----------------------------------	---	---	--	---

Fonte: (SCHORN, 2019)

Ainda segundo (SCHORN, 2019):

A representação de situações reais para exemplificar quando e os motivos de ter a BIA como um novo tipo de conveniência da rotina financeira nas três campanhas está conectada com as tendências da mutualidade, multiplicidade e sincronicidade.

A BIA (BRADESCO, 2022) chegou para auxiliar na execução de tarefas cotidianas pelos clientes e até funcionários, e serviços muitas vezes que antes demandavam mais tempo, um processo de aprendizado maior e as vezes com burocracias, se tornaram mais fáceis de serem feitos através da assistente virtual, atendendo a comandos de voz, textos e muitas vezes com poucos toques na tela do celular. Dessa forma, através dessas interações, como dito anteriormente, a IA consegue se tornar mais eficiente e assertiva em suas respostas, uma vez que através desses comandos e do feedback dos usuários, seu processamento compreende recursos mais específicos, relevantes e de acordo com a vivência de cada um que utiliza o app (SCHORN, 2019).

Um ponto importantíssimo a destacar é o fato do receio que as pessoas têm sobre a substituição humana pela tecnologia, e neste caso, pelas implementações de inteligência artificial, e pode-se observar que de fato a presença delas estão cada vez mais comuns no cotidiano de todos. A dificuldade das pessoas de se adaptarem de forma mais rápida e de maneira flexível em relação a evolução acelerada dessas tecnologias, acaba causando uma certa insegurança. A questão em torno do ambiente que está sendo criado acompanhado dessas interações homem x máquina também tem se tornado um motivo de medo, já que o ambiente se torna mais vertiginoso com o volume e velocidade que as informações são disponibilizadas (SCHORN, 2019).

Atualmente o smartphone é considerado uma parte de nós, e essa união entre o biológico e o tecnológico acaba nos transformando em seres cíbridos. Sobre o conceito dessa palavra, (GABRIEL, 2019) nos diz que:

Cíbridos – híbridos de material e ciberespaço – são entidades que não poderiam existir sem reconciliar a nova classe de símbolos com a materialidade que eles carregam. [...] Cíbridos são mais que simplesmente uma separação completa (entre material e simbólico). Entre esses dois podemos ter componentes compartilhados.

Isso reforça a ideia de que a percepção da dificuldade do ser humano e do meio digital no qual ele está inserido se separarem, tem se tornada cada vez forte. O termo *transhumanism* nos traz um ponto adicional acerca desse assunto, e seu intuito é trazer a tecnologia para mais perto das pessoas para que se beneficiem dela para aumentar suas habilidades nos mais diversos campos. Isso significa que se pode realizar mais tarefas, de forma mais rápida e muitas vezes com mais recursos através da tecnologia, tornando-as cada vez mais simples. Esse conceito *multitasking* nos permite finalizar as demandas e conseguir tempo para aproveitar a família e amigos.

Conforme trata (SCHORN, 2019), pôde-se perceber uma alteração visual na marca do Bradesco nas três campanhas que foram lançadas, principalmente no logotipo e suas formas. O uso elementos que nos lembram o digital, como gráficos de ondas sonoras e linhas conexões virtuais que nos remetem a conexões neurais, nos instigam a entender que se trata de um ambiente totalmente digital, mas que procura se comportar e interagir sempre em busca da forma mais humana possível e aproximar o que de fato estão tentando implementar para o público do setor bancário.

Figura 8 - Logotipos da campanha da BIA e logotipo oficial do Bradesco



Fonte: (SCHORN, 2019)

Na logo da campanha, visualmente alterada para combinar com o lançamento da BIA (BRADESCO, 2022), acrescentando elementos gráficos que ilustra conexões digitais, reforça que o Bradesco deseja estar à frente da concorrência, inovando tecnologicamente e buscando evolução dos seus produtos e forma de atendimento. A Inteligência Artificial foi personificada para se parecer mais humana em suas interações, reforçando também que o banco criou estratégias de humanização e tornou mais conveniente os atendimentos individuais que a empresa fornece. Isso acabou tornando o Bradesco, que sempre foi conhecido por sua tradicionalidade, uma organização mais moderna, atualizada, que traz inovação tecnológica e que deseja transformar o contato com seus clientes uma interação mais rápida, acessível e de fácil comunicação.

(SCHORN, 2019) apresentou 5 hipóteses que serviram de base para conduzir as reflexões sobre o conteúdo exibido e se os objetivos foram atingidos através das estratégias estabelecidas. São elas.

- 1) Ao personificar sua IA, o Bradesco busca humanizar sua relação com os millennials.
- 2) O Bradesco pretende desmistificar que o atendimento bancário é frio, distante e burocrático ao incentivar o contato via IA.
- 3) O uso da IA como canal de atendimento é uma alternativa para sincronizar o Bradesco com o dia a dia dos millennials.
- 4) O Bradesco está em busca de adequar seu posicionamento e sua linguagem via IA para reforçar sua conexão com os millennials.
- 5) O Bradesco está desenvolvendo o senso da transparência e segurança ao falar abertamente sobre a troca de dados e a tecnologia utilizada.

Schorn (2019) conclui sua análise com base nas três campanhas sobre o lançamento da BIA (BRADESCO, 2022) nos mostrando o resultado da validação das hipóteses apresentadas. A Tabela 3 nos mostra que a hipótese 1 está parcialmente validada e que as hipóteses 2, 3 e 4 estão validadas em todas as intenções apresentadas.

Tabela 3 – Cruzamento entre as intenções mapeadas com as hipóteses

Intenção mapeadas do Bradesco	Hipóteses	
Bradesco apresenta a BIA como uma tecnologia inovadora que dá vida à uma nova realidade de relacionamento que já está acessível à população em geral e que tem como objetivo facilitar a rotina financeira das pessoas e também o contato entre clientes e o banco.	Hipótese 1	Validada
	Hipótese 2	Validada
	Hipótese 3	Validada
	Hipótese 4	Validada
	Hipótese 5	Invalidada
Bradesco utiliza a BIA como porta-voz para apresentar-se como um banco que está à frente de seus concorrentes, para destacar seus diferenciais, sendo que um destes é a própria conveniência em ter a BIA como facilitadora da rotina financeira.	Hipótese 1	Invalidada
	Hipótese 2	Validada
	Hipótese 3	Validada
	Hipótese 4	Validada
	Hipótese 5	Invalidada
Bradesco abre o diálogo sobre a relação necessária entre a BIA e os humanos com a intenção de gerar mais interações e empatia com o processo de aprendizado da IA	Hipótese 1	Validada
	Hipótese 2	Validada
	Hipótese 3	Validada
	Hipótese 4	Validada
	Hipótese 5	Validada
Conclusão	Hipótese 1	Parcialmente validada
	Hipótese 2	Completamente validada
	Hipótese 3	Completamente validada
	Hipótese 4	Completamente validada
	Hipótese 5	Invalidada

Fonte: (SCHORN, 2019)

Entende-se, neste caso, que a busca do banco para tornar as interações com a assistente virtual mais humanas vem esse tornando um aliado a própria tecnologia, já que o objetivo é limpar as impressões ruins em relação ao atendimento que os bancos prestam, muitas vezes classificados como frios e de muita burocracia. Dessa forma, a imagem de um atendimento mais assertivo, personalizado e mais humano, se mostra mais favorável e um grande diferencial para a marca no quesito comunicação e comportamento quando uma comunicação entre cliente e banco se faz necessária. A autora termina afirmando que a hipótese 5 foi abordada de forma que seu destaque foi menor em relação as outras, onde foi-se discutido em apenas uma das três campanhas aqui analisadas. Falar sobre a IA com o intuito de trazer mais transparência e mostrar que é seguro utilizá-la acabou tornado essa última hipótese invalidada, já que como foi dito, foi pouco abordada.

Conforme destaca (SCHORN, 2019), o banco busca tornar as interações com a assistente virtual mais humanas vem esse tornando um aliado a própria tecnologia, já que o objetivo é limpar as impressões ruins em relação ao atendimento que os bancos prestam, muitas vezes classificados como frios e de muita burocracia. Dessa forma, a imagem de um atendimento mais assertivo, personalizado e mais humano, se mostra mais favorável e um grande diferencial para a marca no quesito comunicação e comportamento quando uma comunicação entre cliente e banco se faz necessária. Falar sobre a IA, com o intuito de trazer mais transparência e mostrar que é seguro utilizar a ferramenta BIA (BRADESCO, 2022), mostrou-se uma abordagem com pouco resultado relevante.

6.5. O Impacto da BIA no segmento bancário

Atualmente o segmento bancário é um dos mais tradicionais em atuação no Brasil. Nos últimos anos então, tem-se investido fortemente em segurança digital e tecnologia. Por envolver o dinheiro das pessoas, esse tipo de negócio envolve várias frentes de investimento para que todo o sistema seja seguro e eficiente. Estima-se que o lucro líquido médio dos cinco maiores bancos do Brasil - Itaú, Caixa, Bradesco, Santander e Banco do Brasil - seja em torno dos R\$ 13 bilhões conforme dados fornecidos pelo Banco Central (SCHORN, 2019).

Outro ponto importante destacado por (SCHORN, 2019) é que essas grandes empresas, por terem infraestruturas tecnológicas gigantescas, enfrentam desafios cada vez maiores acerca do valor investido e do tempo necessário para manter e atualizar os serviços que disponibilizam. Isso significa que mesmo que as empresas não tenham digitalizado seus processos totalmente,

a era digital tem cobrado a entrega dessas experiências digitais, como por exemplo, na criação de uma conta de forma 100% digital, sem a necessidade do cliente se deslocar até uma agência bancária mais próxima.

Em relação ao uso de smartphones para realizar transações bancárias, "Em 2008, as transações bancárias via smartphone não chegavam a 100 milhões, em 2018, atingiram a marca de 31,3 bilhões, o que representa 40% de todas as transações realizadas no ano, contra 35% em 2017, conforme dados publicados pela Pesquisa FEBRABAN da Tecnologia Bancária 2019.". (SCHORN, 2019)

Os serviços mais utilizados pelos clientes nos smartphones estão a consulta de saldo e extrato, pagamento de boletos, transferências entre contas e contratação de empréstimos. Como muitas pessoas se dirigiam até as agências bancárias para realizar, muitas vezes, esses serviços mais simples, ocorreu um movimento de fechamento de várias agências por todo o país. (SCHORN, 2019).

Em 2017, o Bradesco deu um grande passo em seu crescimento no meio digital e disponibilizou uma ferramenta de Inteligência Artificial, denominada BIA (BRADESCO, 2022), a qual foi desenvolvida em parceria com IBM e utilização da tecnologia Watson. O intuito era principalmente otimizar processos já existentes, o uso de dados dentro da organização e inovar na relação entre humanos e *chatbots*. A IA foi classificada pelos clientes do banco como uma assistente pessoal para ajudar nas transações bancárias, fornecendo dicas de como utilizar o aplicativo, esclarecer dúvidas e operar alguns comandos mais simples. Em 2019 foi atingida a marca de R\$ 100 milhões de interações (SCHORN, 2019).

7. CONCLUSÃO

No presente capítulo, são apresentadas as considerações finais acerca do que foi abordado no decorrer deste trabalho. Inicialmente os objetivos são demonstrados e como eles foram alcançados. Em seguida, são sugeridos alguns pontos que servirão de base para trabalhos futuros.

7.1. Síntese dos Resultados

No presente trabalho foi efetuada uma análise da ferramenta de diálogo automatizado (*chatbot*) denominada BIA, a inteligência artificial do Banco Bradesco, a partir de um levantamento bibliográfico e estudo de caso.

Através do levantamento bibliográfico sobre as tecnologias envolvidas na implementação de ferramentas de diálogo inteligente, foi possível identificar aspectos tecnológicos relevantes como abordagens de implementação, conjuntos de dados utilizados na implementação, os principais métodos de avaliação e a aplicação de softwares de diálogo automatizado (*chatbot*) em diferentes domínios. Também foi possível a identificação dos desafios e limitações nos cenários de aplicação. E por fim, através da análise da BIA (Inteligência Artificial do Bradesco), pôde-se levantar a história de seu surgimento, os desafios encontrados tanto na sua implementação quanto na escolha do conjunto de dados, os impactos na companhia, assim como compreender parte de seu funcionamento por meio do fluxo de telas do uso do aplicativo.

Pôde ser entendido, num contexto geral, o funcionamento dos *chatbots* e sua suma importância no que diz respeito ao atendimento de clientes, análise de informações e direcionamento pertinentes a assuntos específicos quando tratados com o intuito de obter resultados mais assertivos e eficientes com base em experiências anteriores. Casos como os do Banco Bradesco, se mostram como uma página virada rumo ao que se entendia como atendimento humano presencial ou não, e como um atendimento virtual, mesmo que sem a presença de um ser humano diretamente, pode ser rápido, escalável, eficaz e eficiente, confiável e mais autossustentável.

7.2. Resultados e Discussões

7.2.1. Implementação de chatbots

No levantamento bibliográfico sobre as tecnologias envolvidas na implementação de *chatbots* foram consultados os trabalhos de revisão sistemática de literatura (ABDUL-KADER; WOODS, 2015; CAHN, 2017; CALDARINI; JAF; MCGARRY, 2022; MOUSSIADES; ADAMOPOULOU, 2020; NURUZZAMAN; HUSSAIN, 2018; SINGH *et al.*, 2018) e tornou-se possível chegar à análise de que o uso de IA, especificamente *Deep Learning* e PLN, apresentam ser as melhores tecnologias para implementação de chatbots, conforme foi abordado no subcapítulo 4.2. Com o aumento do poder computacional e compartilhamento de tecnologias de código aberto, os algoritmos baseados em recuperação de dados utilizando dessas tecnologias tornaram os chatbots mais fáceis de implementar, mais flexíveis e cada vez mais capazes de imitar a conversação humana. Foi possível chegar à conclusão de que mesmo ainda existindo um grande caminho a ser percorrido, até o momento grandes conquistas foram alcançadas, principalmente devido ao acesso ao grande volume de informações disponíveis na internet. O grande gargalho ainda persiste no quesito conjunto de dados, para que os *chatbots* possam ser mais bem treinados e coerentes. A dificuldade do processamento e aproveitamento desses dados e informações esbarram na falta de um modelo eficaz de avaliação e de treinamento. Ainda sobre o conjunto de dados, conforme destacado na literatura, é possível concluir que quando são escassos, e, de certa forma, ineficientes para o propósito geral. Encontrar uma fonte confiável é uma tarefa cara, demorada e conta com o agravante no âmbito da ética, uma vez que tratar dados mais sensíveis e específicos, que de fato ajudariam a tratar os mais diversos problemas, se torna dispendioso.

Quanto a sua implementação, observa-se que podem ser divididos em dois grupos, sendo chatbots baseados em regras e chatbots de IA. Chatbots baseados em regras, apesar de mais fáceis de implementar, são limitados em termos de capacidade. Uma vez que as regras são altamente específicas de domínio e codificá-las manualmente é difícil e demorado. Já os chatbots de IA utilizando algoritmos de aprendizado de máquina, aprendem através de bases de dados de conversas humanas e também a cada interação que é realizada. Dessa forma, não há a necessidade da implementação de regras, bastando um algoritmo de treinamento o que os torna mais flexíveis e independentes de conhecimento específico.

No tocante a implementação de assistentes virtuais bem treinados, o caminho, apesar de desafiador em vários sentidos, apresenta várias faces satisfatórias e que, a partir de novos estudos direcionados a preencher as lacunas desse processo, pode resultar em *chatbots* ainda mais impressionantes quando a interação homem x máquina acontece. As empresas têm demonstrado grande interesse nessa área pois seu valor e retorno obtido mostram-se bastante encantadores, principalmente por atingir um número maior de pessoas/clientes ao mesmo tempo, independente do lugar e horário e conseguir entregar de forma mais rápida e confiável o que é solicitado, se tornando cada vez melhor à medida que é utilizada.

É importante destacar que se observa um padrão quanto ao modelo de implementação em alguns casos. Quando são aplicados a domínios específicos, os chatbots tendem a recorrer a sistemas de recuperação de informações, isto é, aqueles que buscam respostas em um conjunto pré-definido e/ou baseados em regras.

7.2.2. Aspectos tecnológicos nas implementações

Como apresentado em (CALDARINI; JAF; MCGARRY, 2022), no subcapítulo 4.1, identificando os principais aspectos tecnológicos das ferramentas de diálogo automatizado, no aspecto história, é possível afirmar que houve diversos avanços desde o seu surgimento até os estudos mais atuais. Em 1966, com o surgimento da ELIZA (WEIZENBAUM, 1996), o uso de PLN inspirou vários projetos futuros e os mais diversos estudos na área, o que culminou anos mais tarde no surgimento da ALICE (ATWELL; ABUSHAWAR, 2015), utilizando Inteligência Artificial e XML como base do seu algoritmo. Com isso, foi possível obter uma melhoria significativa em sistemas anteriores de correspondência de padrões e tornando a sua implementação, em termos gerais, mais flexível.

7.2.3. Avaliação das ferramentas

Pode-se concluir também, baseado no que foi apresentado no subcapítulo 4.4, que apesar de serem métricas de avaliação mais pertinentes a tradução automática, as métricas BLEU, METEOR E TER, mencionadas pelos autores (SAIKH *et al.*, 2017; SORDONI *et al.*, 2015), ajudam a avaliar a saída dessas interações. Considera-se a métrica F-score como um meio de dizer se um modelo é compatível dado um conjunto de dados, e sua avaliação categoriza as saídas como positivas ou negativas, ajudando a concluir a qualidade do resultado.

Como apresentado na seção 4.4, avaliar a eficiência e eficácia dos chatbots é considerada ainda algo bastante desafiador e dispendioso. Isso se deve ao fato de que se espelhar na conversa humana envolve vários aspectos, tanto no modo de como chegar e onde chegar nas interações. Dois principais meios de avaliação foram citados: avaliação humana e métricas de avaliação automatizada. Desta forma, várias métricas podem ser aplicadas a vários contextos dependendo do objetivo a ser alcançado, sendo que em alguns casos a avaliação humana também é considerada. Por estes motivos, certas estratégias se tornam caras, demoradas, não escaláveis e abre precedentes para avaliações parciais.

Apesar desses pontos negativos, o nível de crítica e amplitude de aspectos que podem ser alcançados são importantes e são retratados em várias partes das literaturas analisadas, onde podem ser citados alguns dos quadros de avaliação mais utilizados, como o *PARAdigm for Dialogue System Evaluation*.

7.2.4. Áreas de aplicação e interesses

No que se refere a aplicação dos chatbots, múltiplos domínios podem usufruir da tecnologia. Áreas da Educação e Pesquisa, assim como da Saúde, são demonstrados pelos autores que possuem diversos exemplos de aplicações e que se baseiam principalmente nos modelos de implementação com recuperação de informações, conforme exposto no subcapítulo 4.5. Algumas outras áreas que também utilizam esse modelo de implementação principalmente, podem ser citadas, como de Finanças e Treinamentos. Já, como citado por (JOHN *et al.*, 2017; SHEIKH; TIWARI; SINGHAL, 2019), os modelos de sequência a sequência também são vistos no campo de Direitos e RH, baseados em LSTM e Unidades Recorrentes com Porta Bidirecional, respectivamente.

É possível concluir também que a confiança das informações no fluxo das interações é algo extremamente importante, e reforça a ideia de que ainda é necessária uma métrica de avaliação que considere vários pontos que contribuem para resultados cada vez mais confiáveis, como eficiência, coesão e qualidade dos diálogos. Ainda pode-se concluir que apesar das métricas de avaliação automatizadas serem mais eficientes em termos de tempo e recursos, ainda carece de um padrão e da competência de avaliar corretamente o diálogo como um todo. Mesmo assim, são implementadas de forma majoritária por sua facilidade de usar.

7.2.5. Estudo de caso: BIA

Ao falar sobre a Inteligência Artificial BIA do banco Bradesco, como apresentado no Capítulo 6, constata-se que foi uma grande revolução no âmbito bancário pela forma de atendimento nada tradicional e ambiciosa. Apesar das grandes dificuldades apontadas na implementação dos *chatbots*, a empresa tem colhido grandes resultados, com destaque para a qualidade e agilidade nos atendimentos aos clientes e para os colaboradores, sendo muitos deles treinados pela própria IA a como executar suas funções, operações e tirar dúvidas acerca de informações mais sensíveis, como taxas e valores.

A BIA (BRADESCO, 2022) foi desenvolvida para oferecer atendimento imediato as respostas dos clientes. Criada a partir da plataforma de computação cognitiva Watson em parceria com a IBM, a inteligência artificial funciona com base no comportamento do usuário. A produção do sistema levou cerca de dois anos. Primeiro ficou restrita ao uso interno, por gerentes do banco, e depois foi apresentada ao grande público, ou seja, aos milhões de correntistas do Bradesco. Hoje além de tirar dúvidas também é capaz de realizar transações bancárias.

A assistente virtual do aplicativo Bradesco funciona por meio da computação cognitiva, de forma muito simples e intuitiva. Assim como as assistentes do Android, iOS e Microsoft, para interagir com a BIA (BRADESCO, 2022), basta que o usuário faça perguntas e forneça informações sobre sua vida financeira ou pessoal. Por meio da computação cognitiva, o usuário interage com o *chatbot* para que ela consiga aprender quais são suas necessidades e consiga até mesmo antecipá-las. O processo de aprendizado e interação acontece em tempo real e de forma mútua.

No quesito implementação da BIA (BRADESCO, 2022), pode-se concluir que o *chatbot* utiliza o modelo baseado em regras, e que sua evolução nas interações com qualidade se dá conforme a ferramenta vai sendo utilizada. Sua personalidade é neutra e conforme (PAIVA, 2018) futuramente a IA adotará um comportamento diferente para cada cliente de acordo com o contexto e necessidade conforme os diálogos acontecem.

Foi descoberto também que a BIA (BRADESCO, 2022) nasceu de uma parceria do Bradesco com a IBM e que o processamento dos dados acontece nos servidores da própria IBM. Esse conjunto de dados que vêm das interações colabora muito com o crescimento da assistente virtual, mas ao mesmo tempo limita a coerência e eficiência dos resultados uma vez que esse tipo de treinamento através de um conjunto específico não traz riqueza de detalhes e possuem

muitos ruídos, que são informações as vezes sem contexto, com escritas erradas e que também podem ser dados sensíveis.

Foi possível entender que quanto à segurança, a BIA (BRADESCO, 2022), ainda necessita de estudos mais aprofundados acerca da melhor maneira de lidar com o tráfego de dados privados e que mesmo entregando informações incompletas é possível obter bons resultados (PAIVA, 2018).

Para concluir sobre a implementação da BIA (BRADESCO, 2022), um ponto importante a destacar é do feedback. Sempre que uma interação ocorre, o usuário é capaz de avaliar se o resultado é coerente com o que foi solicitado e se a informação ajudou o cliente naquele contexto. Isso, como já foi dito, ajuda no conhecimento da ferramenta e a torna cada vez mais rápida e eficiente.

Conclui-se também que o objetivo geral da implementação e uso dos chatbots, especialmente no Bradesco, se mostrou bastante promissor e que os objetivos foram atingidos. A solução de questões de forma simples, disponível a todo o momento e com a escalabilidade de alcançar um número maior de clientes foi destaque e colocou o Bradesco na posição de primeiro banco ao lançar um assistente virtual com uso de IA com milhões de interações e de fácil uso.

A partir da análise das simulações de uso do aplicativo da BIA (BRADESCO, 2022) pode-se concluir que seu uso é bastante fluido intuitivo. A IA, através de regras linguísticas e técnicas de correspondência de padrões identifica o que o cliente está pedindo e retorna com uma resposta pré-definida baseada no que foi solicitado. As informações repassadas têm “níveis” diferentes dependendo do tipo do acesso, se acessada a conta ou não, o que mostra preocupação quanto à segurança e proteção dos dados dos usuários. Transações bancárias mais simples também podem ser feitas de forma bastante simples, apenas informando os dados através do teclado do dispositivo ou por comando de voz através do microfone. Sempre que um diálogo chega ao fim é solicitado um feedback ao usuário se a resposta foi satisfatória ou não, e com este feedback a IA “aprende” quais as saídas mais pertinentes para outras perguntas do mesmo tipo de outros usuários.

7.3. Limitações e Trabalhos Futuros

Pelo detalhamento apresentado em (CALDARINI; JAF; MCGARRY, 2022) no Capítulo 5, observando os desafios, limitações nos cenários de aplicação de *chatbots*, pode-se

afirmar que, apesar da ideia inicial dessa tecnologia ter surgido em 1950, ainda hoje se faz necessário aprofundar-nos no tema já que ainda não foi determinado um modelo de implementação e avaliação das ferramentas de forma global. Ainda pode-se acrescentar que o problema de se conseguir um conjunto de dados amplo, sem ruídos e para todas as áreas de aplicação, perdura e torna difícil a propagação e uso extenso dos chatbots nas mais diversas aplicações específicas da indústria. Mesmo com os avanços tecnológicos, o objetivo dos *chatbots* de IA de imitar a fala humana ainda não foi alcançado e o caminho a ser percorrido para chegar a uma interação quase que idêntica a com um humano é grande.

Sobre os conjuntos de dados para treinamento dos chatbots de IA observa-se que na literatura pesquisada não há conjunto de dados de domínio fechado usados com frequência provavelmente devido a sua especificidade de domínio. Quanto ao conjunto de dados de domínio aberto pode-se concluir que possuem muito ruído quando são volumosos, já quando os dados são mais úteis e limpos também são escassos, portanto, menos adequados para treinamento de diálogos mais complexos.

Apesar dos resultados terem mostrado a evolução desde o século passado acerca do desenvolvimento de chatbots e sua expansão das áreas de aplicação, estudos envolvendo um padrão de implementação e de um modelo de avaliação mais eficiente seria interessante como forma de obter interações cada vez mais naturais e parecidas com diálogos humanos. Um outro ponto que poderia ser considerado é a possibilidade de substituir a avaliação humana de forma inteligente e com todas as qualidades pertinentes, já que são caras, demoradas e possuem vieses. Será mesmo possível implementar tais modelos de avaliação e métricas e aumentar a escalabilidade dessas avaliações?

Pode-se sugerir para trabalhos futuros o estudo de novos modelos de avaliação mais eficientes e escaláveis, bem como formas de obter conjuntos de dados para treinamento dos *chatbots* cada vez mais ricos em dados e de qualidade, sem ferir a ética do processo. Por fim, sugestão de aplicações em cada vez mais áreas podem ser importantes, visto que seu uso tem devolvido grandes resultados à companhia como um todo.

Ainda, novos modelos de implementação e conjunto de dados para treinamento se mostram de uma importância imprescindível, uma vez que os *chatbots* precisam ser bem treinados para identificar padrões de comportamentos incorretos que podem levar a falsos positivos baseados no aprendizado de máquina. Também quanto à acessibilidade é importante que seja estudada a capacidade dos *chatbots* de compreender e lidar com diferentes idiomas ou

peessoas com as mais diversas deficiências para que sejam sempre assertivos e respeitosos em todo e qualquer caso.

No que diz respeito aos assédios que acontecem nas interações com os *chatbots*, algumas soluções apresentadas se referem ao treinamento das aplicações para que sejam capazes de responder as mais diversas formas de violência. Além disso, investir em educação sobre o tema pode ser tornar fundamental para a mudança desse cenário. É um tema importante que vale muito a pena ser estudado, seja em formas de programar e treinar esses assistentes virtuais para lidar com um número máximo de possibilidade, e em programas de incentivo a educação e conscientização acerca do assunto.

REFERÊNCIAS BIBLIOGRÁFICAS

ABDUL-KADER, Sameera A.; WOODS, John C. Survey on chatbot design techniques in speech conversation systems. **International Journal of Advanced Computer Science and Applications**, v. 6, n. 7, 2015.

ADIWARDANA, Daniel *et al.* Towards a human-like open-domain chatbot. **arXiv preprint arXiv:2001.09977**, 2020.

AGARWAL, Abhaya; LAVIE, Alon. Meteor, m-bleu and m-ter: Evaluation metrics for high-correlation with human rankings of machine translation output. *In: PROCEEDINGS OF THE THIRD WORKSHOP ON STATISTICAL MACHINE TRANSLATION*. 2008. 2008, p. 115-118.

ALASUNDRAM, Kulothunkan *et al.* Sequence to sequence model performance for education chatbot. **International journal of emerging Technologies in Learning (iJET)**, v. 14, n. 24, p. 56-68, 2019.

ALI, Maha Mahmoud; KUMAR, Ramakrishna. A review on chatbot design and implementation techniques. **Int. J. Eng. Technol**, v. 7, n. 11, 2020.

ARIFI, Jeton; EBNER, Markus; EBNER, Martin. Potentials of Chatbots for Spell Check among Youngsters. **Int. J. Learn. Anal. Artif. Intell. Educ.**, v. 1, n. 1, p. 54-60, 2019.

ATHOTA, Lekha *et al.* Chatbot for healthcare system using artificial intelligence. *In: 2020 8TH INTERNATIONAL CONFERENCE ON RELIABILITY, INFOCOM TECHNOLOGIES AND OPTIMIZATION (TRENDS AND FUTURE DIRECTIONS) (ICRITO)*. IEEE, 2020, p. 619-622.

ATWELL, Eric; ABUSHAWAR, Bayan. ALICE chatbot: Trials and outputs. **Computación y Sistemas**, v. 19, n. 4, p. 625-632, 2015.

AUGELLO, Agnese *et al.* An approach to enhance chatbot semantic power and maintainability: experiences within the FRASI project. *In: 2012 IEEE SIXTH INTERNATIONAL CONFERENCE ON SEMANTIC COMPUTING*. IEEE, 2012, p. 186-193.

AYANOUZ, Soufyane; ABDELHAKIM, Boudhir Anouar; BENHMED, Mohammed. A smart chatbot architecture-based NLP and machine learning for health care assistance. *In: PROCEEDINGS OF THE 3RD INTERNATIONAL CONFERENCE ON NETWORKING, INFORMATION SYSTEMS & SECURITY*. 2020, p. 1-6.

BAHDANAU, Dzmitry; CHO, Kyunghyun; BENGIO, Yoshua. Neural machine translation by jointly learning to align and translate. **arXiv preprint arXiv:1409.0473**, 2014.

BALA, K *et al.* **Chat-Bot for College Management System Using A.I.** 2017. Disponível em: <https://www.irjet.net/archives/V4/i11/IRJET-V4I11367.pdf>. Acesso em: 13 out. 2022.

BANERJEE, Satanjeev; LAVIE, Alon. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. *In: PROCEEDINGS OF THE ACL WORKSHOP ON INTRINSIC AND EXTRINSIC EVALUATION MEASURES FOR MACHINE TRANSLATION AND/OR SUMMARIZATION*. 2005. 2005. 8 p.

BENTON, Morgan C.; RADZIWIŁŁ, Nicole M. Evaluating quality of chatbots and intelligent conversational agents. **arXiv preprint arXiv:1704.04579**, 2017.

BERGER, Rene; EBNER, Markus; EBNER, Martin. Conception of a Conversational Interface to Provide a Guided Search of Study Related Data. **International Journal of Emerging Technologies in Learning**, v. 14, n. 7, 2019.

BERNARDINI, Andréia Ana; SÔNEGO, Arildo A.; POZZEBON, Eliane. Chatbots: An analysis of the state of art of literature. *In: ANAIS DO I WORKSHOP ON ADVANCED VIRTUAL ENVIRONMENTS AND EDUCATION*. SBC, 2018. 1 p.

BIA [sistema eletrônico]. Disponível em: <<https://banco.bradesco/canaisdigitais/conheca-bia.shtm>>. Acesso em: 10 dez. 2022.

BRANDTZAEG, Petter Bae; FØLSTAD, Asbjørn. Why people use chatbots. *In: INTERNATIONAL CONFERENCE ON INTERNET SCIENCE*. 2017, p. 377-392.

CAHN, Jack. CHATBOT: Architecture, design, & development. **University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science**, 2017.

CALDARINI, Guendalina; JAF, Sardar; MCGARRY, Kenneth. A literature survey of recent advances in chatbots. **Information**, v. 13, n. 1. 41 p, 2022.

CASILLO, Mario *et al.* Chatbot in industry 4.0: An approach for training new employees. *In: 2020 IEEE INTERNATIONAL CONFERENCE ON TEACHING, ASSESSMENT, AND LEARNING FOR ENGINEERING (TALE)*. IEEE, 2020, p. 371-376.

CHEN, Stanley F.; BEEFERMAN, Douglas; ROSENFELD, Roni. **Evaluation metrics for language models**. 1998.

CHRISTENSEN, Silje *et al.* Context-aware sequence-to-sequence models for conversational systems. **arXiv preprint arXiv:1805.08455**, 2018.

COSTA, Pedro. Conversing with personal digital assistants: on gender and artificial intelligence. **Journal of Science and Technology of the Arts**, v. 10, n. 3, p. 59-72, 2018.

CUAYÁHUITL, Heriberto *et al.* Ensemble-based deep reinforcement learning for chatbots. **Neurocomputing**, v. 366, p. 118-130, 2019.

CUI, Lei *et al.* Superagent: A customer service chatbot for e-commerce websites. *In: PROCEEDINGS OF ACL 2017, SYSTEM DEMONSTRATIONS*. 2017, p. 97-102.

DAI, Zihang *et al.* Transformer-xl: Attentive language models beyond a fixed-length context. **arXiv preprint arXiv:1901.02860**, 2019.

DANESCU-NICULESCU-MIZIL, Cristian; LEE, Lillian. Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. **arXiv preprint arXiv:1106.3077**, 2011.

DAUPHIN, Yann N. *et al.* Language modeling with gated convolutional networks. *In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING*. PMLR, 2017, p. 933-941.

DEVLIN, Jacob *et al.* Bert: Pre-training of deep bidirectional transformers for language understanding. **arXiv preprint arXiv:1810.04805**, 2018.

DHYANI, Manyu; KUMAR, Rajiv. An intelligent Chatbot using deep learning with Bidirectional RNN and attention model. **Materials today: proceedings**, v. 34, p. 817-824, 2021.

DORMEHL, Luke. Microsoft's Friendly XiaoIce AI Can Figure Out What You Want—Before You Ask. **DigitalTrends (November 18)**, www.digitaltrends.com/cool-tech/xiaoice-microsoft-future-of-ai-assistants, 2018.

FERNANDES, Anush. **NLP, NLU, NLG and how Chatbots work**. Chatbots Life. 2017. Disponível em: <https://chatbotslife.com/nlp-nlu-nlg-and-how-chatbots-work-dd7861dfc9df>. Acesso em: 20 jul. 2022.

GABRIEL, M. CIBRIDISMO: APRENDA O CONCEITO POR TRÁS DO TEMA. Gen. Negócios e Gestão, 2019. Disponível em: < <https://gennegociosegestao.com.br/cibridismo-aprenda-conceito//>> Acesso em: 07 dez. 2022.

GHANDEHARIOUN, Asma *et al.* Approximating interactive human evaluation with self-play for open-domain dialog system. **Advances in Neural Information Processing Systems**, v. 32, 2019.

GOMES, SANDRA LÚCIA RÉBEL; MENDONÇA, MARÍLIA ALVARENGA ROCHA; SOUZA, Clarice Muhlethaler de. Literatura cinzenta. *In: Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: Ed. UFMG, 2000. p. 97-104.

HE, Tianxing *et al.* Analyzing the forgetting problem in the pretrain-finetuning of dialogue response model. **arXiv preprint arXiv:1910.07117**, 2019.

HIGH, Rob. The era of cognitive systems: An inside look at IBM Watson and how it works. **IBM Corporation, Redbooks**, v. 1, p. 16, 2012.

HIGASHINAKA, Ryuichiro *et al.* Towards an open-domain conversational system fully based on natural language processing. *In: PROCEEDINGS OF COLING 2014, THE 25TH INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS: TECHNICAL PAPERS*. 2014. 2014. 12 p.

HUNG, Victor *et al.* Towards a method for evaluating naturalness in conversational dialog systems. *In: 2009 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS*. IEEE, 2009, p. 1236-1241.

JIA, Jiyou. **The study of the application of a keywords-based chatbot system on the teaching of foreign languages**. arXiv preprint cs/0310018. 2013. Disponível em: <https://arxiv.org/abs/cs/0310018>. Acesso em: 2 nov. 2022.

JOHN, Adebayo *et al.* Legalbot: A deep learning-based conversational agent in the legal domain. *In: INTERNATIONAL CONFERENCE ON APPLICATIONS OF NATURAL LANGUAGE TO INFORMATION SYSTEMS*. Springer, Cham, 2017. 273 p.

JURAFSKY, Dan; MARTIN, James H. **Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition**. Prentice Hall, f. 514, 2008. 1027 p.

KANNAN, Anjuli; VINYALS, Oriol. Adversarial evaluation of dialogue models. **arXiv preprint arXiv:1701.08198**, 2017.

KEMPSON, Ruth; GARGETT, Andrew; GREGOROMICHELAKI, Eleni. Clarification requests: An incremental account. **Decalog 2007**. 65 p, 2007.

KIM, Jintae *et al.* Multi-turn chatbot based on query-context attentions and dual wasserstein generative adversarial networks. **Applied Sciences**, v. 9, n. 18. 3908 p, 2019.

KITAEV, Nikita; KAISER, Łukasz; LEVSKAYA, Anselm. Reformer: The efficient transformer. **arXiv preprint arXiv:2001.04451**, 2020.

KLEIN, Guillaume *et al.* Opennmt: Open-source toolkit for neural machine translation. **arXiv preprint arXiv:1701.02810**, p. 67-72, 2017.

KUKSENOK, Kit; MARTYNIV, Andriy. Evaluation and improvement of chatbot text classification data quality using plausible negative examples. **arXiv preprint arXiv:1906.0191**, p. 87-95, 2019.

LEAL, Renato. **Por que chatbots falham? E como evitar que isso aconteça?** LinkedIn, 2018. Disponível em: < <https://pt.linkedin.com/pulse/por-que-chatbots-falham-e-como-evitar-isso-aconte%C3%A7a-renato-leal> >. Acesso em: 12/02/2023

LI, Di; HE, Xiao-dong; SHUM, Heung-Yeung. From Eliza to XiaoIce: challenges and opportunities with social chatbots. **Frontiers of Information Technology & Electronic Engineering**, v. 19, n. 1, p. 10-26, 2018.

LI, Jiwei *et al.* Adversarial learning for neural dialogue generation. **arXiv preprint arXiv:1701.06547**, 2017.

LI, Jiwei *et al.* Deep reinforcement learning for dialogue generation. **arXiv preprint arXiv:1606.01541**, 2016.

LI, Margaret; WESTON, Jason; ROLLER, Stephen. Acute-eval: Improved dialogue evaluation with optimized questions and multi-turn comparisons. **arXiv preprint arXiv:1909.03087**, 2019.

LI, Yanran *et al.* Dailydialog: A manually labelled multi-turn dialogue dataset. **arXiv preprint arXiv:1710.03957**, p. 986-995, 2017.

LIN, Zhaojiang *et al.* XPersona: Evaluating multilingual personalized chatbot. **arXiv preprint arXiv:2003.07568**, 2020.

LISON, Pierre; TIEDEMANN, Jörg. **Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles**. 2016.

LU, Zhengdong; LI, Hang. A deep architecture for matching short texts. **Advances in neural information processing systems**, v. 26, 2013.

LUO, Xueming *et al.* Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. **Marketing Science**, v. 38, n. 6, p. 937-947, 2019.

MELO, Glauca *et al.* Exploring context-aware conversational agents in software development. **arXiv preprint arXiv:2006.02370**, 2020.

MIRHET, Saracevic; EBNER, Markus; EBNER, Martin. Potential of Bots for Encyclopedia. **IPSI BgD Transactions on Internet Research**, v. 16, n. 1, p. 54-60, 2020.

MLADENIĆ, Dunja; BRADEŠKO, Luka. A survey of chatbot systems through a loebner prize competition. *In: PROCEEDINGS OF SLOVENIAN LANGUAGE TECHNOLOGIES SOCIETY EIGHTH CONFERENCE OF LANGUAGE TECHNOLOGIES*, Ljubljana, Slovenia: Institut Jožef Stefan, 2012, p. 34-37.

MOUSSIADES, Lefteris; ADAMOPOULOU, Eleni. An overview of chatbot technology. *In: IFIP INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE APPLICATIONS AND INNOVATIONS*. 2020, p. 373-383

Os impactos do assédio contra chatbots e assistentes virtuais. NeoAssist, 2020. Disponível em: < https://www.neoassist.com/blog/os-impactos-do-assedio-contra-chatbots-e-assistentes-virtuais/#Vies_inconsciente_qual_a_solucao_para_os_chatbots>. Acesso em: 10 fev. 2023.

NIMAVAT, Ketakee; CHAMPANERIA, Tushar. Chatbots: An overview types, architecture, tools and future possibilities. **Int. J. Sci. Res. Dev**, v. 5, n. 7, p. 1019-1024, 2017.

NOVAS RESPOSTAS DA BIA CONTRA O ASSÉDIO. Bradesco, 2022. Disponível em: <<https://banco.bradesco/aliadosbia/#:~:text=Desde%202018%2C%20quando%20a%20BIA,fe mininos%20e%20tamb%C3%A9m%20sofre%20ass%C3%A9dio>>. Acesso em: 12 fev. 2023.

NURUZZAMAN, Mohammad; HUSSAIN, Omar Khadeer. A survey on chatbot implementation in customer service industry through deep neural networks. *In: 2018 IEEE 15TH INTERNATIONAL CONFERENCE ON E-BUSINESS ENGINEERING (ICEBE)*. 2018, p. 54-61.

NWANKWO, Wilson. Interactive advising with bots: improving academic excellence in educational establishments. **Wilson Nwankwo. Interactive Advising with Bots: Improving Academic Excellence in Educational Establishments. American Journal of Operations Management and Information Systems**, v. 3, n. 1, p. 6-21, 2018.

OIKUMORO, Abayomi Oluwaseyi; JAWAD, Mohammed Saeed. Intention to use intelligent conversational agents in e-commerce among Malaysian SMEs: an integrated conceptual framework based on tri-theories including unified theory of acceptance, use of technology (UTAUT), and TOE. **International Journal of Academic Research in Business and Social Sciences**, v. 9, n. 11, p. 205-235, 2019.

OKUDA, Takuma; SHODA, Sanae. AI-based chatbot service for financial industry. **Fujitsu Scientific and Technical Journal**, v. 54, n. 2, p. 4-8, 2018.

PAIVA, F. A história do nascimento da BIA. Mobile Site, 2018. Disponível em: <<https://www.mobiletime.com.br/noticias/04/05/2018/a-historia-do-nascimento-da-BIA/>> Acesso em: 27 nov. 2022.

PRZEGALINSKA, Aleksandra *et al.* In bot we trust: A new methodology of chatbot performance measures. **Business Horizons**, v. 62, n. 6, p. 785-797, 2019.

RAHMAN, A. M.; AL MAMUN, Abdullah; ISLAM, Alma. Programming challenges of chatbot: Current and future prospective. In: **2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)**. IEEE, 2017. p. 75-78.

RAMACHANDRAN, Prajit; ZOPH, Barret; LE, Quoc V. Searching for activation functions. **arXiv preprint arXiv:1710.05941**, 2017.

ROCKCONTENT. **Quem é a BIA? Conheça a Inteligência Artificial do Bradesco**, 2018. Disponível em: <https://rockcontent.com/br/blog/BIA-bradesco/>. Acesso em: 20 jul. 2022.

ROLLER, Stephen *et al.* Recipes for building an open domain chatbot. **arXiv preprint arXiv:2004.13637**, 2020.

SAIKH, Tanik *et al.* Textual entailment using machine translation evaluation metrics. In: **INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS AND INTELLIGENT TEXT PROCESSING**. Springer, Cham, 2017, p. 317-328.

SASAKI, Yutaka. The truth of the F-measure. **Teach tutor mater**, v. 1, n. 5, p. 1-5, 2007.

SCHORN, Bruna. **O IMPACTO DA INTELIGÊNCIA ARTIFICIAL NA RELAÇÃO DOS BANCOS COM OS MILLENNIALS NA ERA PÓS-DIGITAL: UM ESTUDO DE CASO DA BIA**. Orientador: Eduardo Zilles Borba. 2019. 135 f. TCC (Graduação) – Curso de Comunicação Social – Habilitação em Publicidade e Propaganda, Unidade Acadêmica de Graduação, Universidade Do Vale Do Rio Dos Sinos, São Leopoldo. 2019. Disponível em: <http://www.repositorio.jesuita.org.br/handle/UNISINOS/10970>. Acesso em: 28 nov. 2022.

SEDOC, Joao *et al.* Chateval: A tool for chatbot evaluation. *In: PROCEEDINGS OF THE 2019 CONFERENCE OF THE NORTH AMERICAN CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (DEMONSTRATIONS)*. 2019. 2019, p. 62-65.

SHANG, Lifeng; LU, Zhengdong; LI, Hang. Neural responding machine for short-text conversation. **arXiv preprint arXiv:1503.02364**, 2015.

SHAWAR, Bayan Abu; ATWELL, Eric Steven. Using corpora in machine-learning chatbot systems. *International journal of corpus linguistics*, v. 10, n. 4, p. 489-516, 2005.

SHEEHAN, Ben; JIN, Hyun Seung; GOTTLIEB, Udo. Customer service chatbots: Anthropomorphism and adoption. **Journal of Business Research**, v. 115, n. 14-24, 2020.

SHEIKH, Salim Akhtar; TIWARI, Vineeta; SINGHAL, Sunita. Generative model chatbot for human resource using deep learning. *In: 2019 INTERNATIONAL CONFERENCE ON DATA SCIENCE AND ENGINEERING (ICDSE)*. IEEE, 2019, p. 126-132.

SINGH, Rupesh *et al.* Chatbot using TensorFlow for small Businesses. *In: 2018 SECOND INTERNATIONAL CONFERENCE ON INVENTIVE COMMUNICATION AND COMPUTATIONAL TECHNOLOGIES (ICICCT)*. IEEE, 2018, p. 1614-1619.

SNOVER, Matthew *et al.* A study of translation edit rate with targeted human annotation. *In: PROCEEDINGS OF THE 7TH CONFERENCE OF THE ASSOCIATION FOR MACHINE TRANSLATION IN THE AMERICAS: TECHNICAL PAPERS*. 2006, p. 223-231.

SO, David; LE, Quoc; LIANG, Chen. The evolved transformer. *In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING*. PMLR, 2019, p. 5877-5886.

SOJASINGARAYAR, Abonia. **Seq2seq ai chatbot with attention mechanism**. arXiv preprint arXiv:2006.02767. 2020. Disponível em: <https://arxiv.org/abs/2006.02767>. Acesso em: 27 set. 2022.

SORDONI, Alessandro *et al.* A neural network approach to context-sensitive generation of conversational responses. **arXiv preprint arXiv:1506.06714**, 2015.

SPENCER, Geoff. Much more than a chatbot: China's Xiaoice mixes AI with emotions and wins over millions of fans. **Microsoft Asia News Cent**, 2018.

STRIGÉR, Amanda. **End-to-end trainable chatbot for restaurant recommendations**. 2017.

SUNDAR, S. Shyam; GO, Eun. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. **Computers in Human Behavior**, v. 57, p. 304-316, 2019.

SUTSKEVER, Ilya; VINYALS, Oriol; LE, Quoc V. Sequence to sequence learning with neural networks. **Advances in neural information processing systems**, v. 27, 2014.

TURING, Alan M. Computing machinery and intelligence. *In: PARSING THE TURING TEST*. Springer, Dordrecht, 2009, p. 23-65.

VASWANI, Ashish *et al.* Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017.

VINYALS, Oriol; LE, Quoc. A neural conversational model. **arXiv preprint arXiv:1506.05869**, 2015.

WALKER, Marilyn *et al.* PARADISE: A framework for evaluating spoken dialogue agents. **arXiv preprint cmp-lg/9704004**, p. 271-280, 1997.

WANG, Yi Fei; PETRINA, Stephen. Using learning analytics to understand the design of an intelligent language tutor–Chatbot lucy. **Editorial Preface**, v. 4, n. 11, p. 124-131, 2013.

WEIZENBAUM, Joseph. ELIZA—a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, v. 9, n. 1, p. 36-45, 1996.

WILCOX, Sue; WILCOX, Bruce. Winning the Loebner's. **Brillig Understanding Inc**, 2014.

XU, Jing *et al.* Recipes for safety in open domain chatbots. **arXiv preprint arXiv:2010.07079**, 2020.

XU, Kelvin *et al.* Show, attend and tell: Neural image caption generation with visual attention. *In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING*. PMLR, 2015, p. 2048-2057.

YAN, Rui; SONG, Yiping; WU, Hua. Learning to respond with deep neural networks for retrieval-based human-computer conversation system. *In: PROCEEDINGS OF THE 39TH INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL*. 2016, p. 55-64.

ZEMČÍK, M. Tmáš. A brief history of chatbots. **DEStech Transactions on Computer Science and Engineering**, v. 10, 2019.

ZHONG, Peixiang; WANG, Di; MIAO, Chunyan. An affect-rich neural conversational model with BIA sed attention and weighted cross-entropy loss. *In: PROCEEDINGS OF THE AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE*. 2019, p. 7492-7500.

ZHOU, Li *et al.* The design and implementation of xiaoice, an empathetic social chatbot. **Computational Linguistics**, v. 46, n. 1, p. 53-93, 2020.